AFRL-OSR-VA-TR-2015-0208

Robust, Adaptive Radar Detection and Estimation

Vishal Monga
PENNSYLVANIA STATE UNIVERSITY THE

07/21/2015
Final Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ RTC
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
14-07-2015	Final Report	07/15/2012 - 07/14/2015

4. TITLE AND SUBTITLE
Robust, Adaptive Radar Detection and Estimation

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-12-1-0333

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Monga, Vishal

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Pennsylvania State University, University Park, PA 16802

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/JA (703) 696-9705
875 North Randolph Street, Suite 325, Room 3112m Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A - Approved for Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This work introduces estimation of disturbance covariance matrices for radar STAP. In particular, we first exploit physically inspired constraints, both the structure of the disturbance covariance and importantly the knowledge of the clutter rank to yield a new rank constrained maximum likelihood (RCML) estimator of clutter/disturbance covariance. We demonstrate that the rank-constrained estimation problem can in fact be cast in the framework of a tractable convex optimization problem, and derive closed form expressions for the estimated covariance matrix. On top of that, this work also introduces a new computationally efficient covariance estimator which jointly enforces a Toeplitz structure and a rank constraint. Our proposed solution focuses on a computationally efficient approximation and involves a cascade of two closed form solutions, the RCML estimator and the rank preserving Toeplitz approximation. Finally, we develop robust estimators that can adapt to imperfect knowledge of physical constraints using an expected likelihood (EL) approach. We analyze covariance estimation algorithms under three cases of imperfect constraints: only rank constraint, both rank and noise power constraints, and condition number constraint.

15. SUBJECT TERMS
adaptive signal processing, convex optimization, estimation

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
U	U	U	UU	87	Vishal Monga
					19b. TELEPHONE NUMBER *(Include area code)* 814-865-7667

Reset

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

OVERALL PROGRESS SUMMARY

- Contract/Grant Title: Robust, Adaptive Radar Detection and Estimation
- Contract/Grant #: FA9550-12-1-0333
- Reporting Period: 15 July 2012 - 14 July 2015
- Annual accomplishments (200 words max):
  This work introduces estimation of disturbance covariance matrices for radar STAP. In particular, we first exploit physically inspired constraints, both the structure of the disturbance covariance and importantly the knowledge of the clutter rank to yield a new rank constrained maximum likelihood (RCML) estimator of clutter/disturbance covariance. We demonstrate that the rank-constrained estimation problem can in fact be cast in the framework of a tractable convex optimization problem, and derive closed form expressions for the estimated covariance matrix. On top of that, this work also introduces a new computationally efficient covariance estimator which jointly enforces a Toeplitz structure and a rank constraint. Our proposed solution focuses on a computationally efficient approximation and involves a cascade of two closed form solutions, the RCML estimator and the rank preserving Toeplitz approximation. Finally, we develop robust estimators that can adapt to imperfect knowledge of physical constraints using an expected likelihood (EL) approach. We analyze covariance estimation algorithms under three cases of imperfect constraints: only rank constraint, both rank and noise power constraints, and condition number constraint. Extensive performance evaluation on both simulated and KASSPER data confirms the proposed estimators excel previous works.
- Archival publications (published) during reporting period:
  1. V. Monga and M. Rangaswamy, "Rank Constrained ML Estimation of Structured Covariance Matrices with Applications in Radar Target Detection," 2012 IEEE Radar Conference
  2. B. Kang, V. Monga, and M. Rangaswamy, "On the Practical Merits of Rank Constrained ML Estimator of Structured Covariance Matrices," 2013 IEEE Radar Conference
  3. B. Kang, V. Monga, and M. Rangaswamy, "Efficient Approximation of Structured Covariance under Joint Toeplitz and Rank Constraints," IEEE Asilomar Conference on Signal, Systems and Computers, November 2013
  4. B. Kang, V. Monga, and M. Rangaswamy, "Constrained ML Estimation of Structured Covariance Matrices with Applications in Radar STAP," Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), The fifth IEEE International Workshop, December 2013
  5. B. Kang, V. Monga, and M. Rangaswamy, "Rank Constrained ML Estimation of Structured Covariance Matrices," IEEE Transactions on Aerospace and Electronic Systems, January 2014
  6. B. Kang, V. Monga, and M. Rangaswamy, "Estimation of Structured Covariance Matrices for Radar STAP under Practical Constraints," 2014 IEEE Radar Conference, **The Best Student Paper Award**
  7. B. Kang, V. Monga, and M. Rangaswamy, "Computationally Efficient Toeplitz Approximation of Structured Covariance under a Rank Constraint," IEEE Transactions on Aerospace and Electronic Systems, January 2015
  8. B. Kang, "Student Research Highlight: Estimation of Structured Covariance Matrices for Radar STAP," IEEE Aerospace and Electronic Systems Magazine, February 2015
  9. B. Kang, V. Monga, M. Rangaswamy, and Y. Abramovich, "Automatic Rank Estimation for Practical STAP Covariance Estimation via an Expected Likelihood Approach," 2015 IEEE Radar Conference
  10. B. Kang, V. Monga, M. Rangaswamy, and Y. Abramovich, "Robust Covariance Estimation under Imperfect Constraints using Expected Likelihood Approach," to be submitted to IEEE Transactions on Aerospace and Electronic Systems

- Changes in research objectives, if any: None
- Change in AFOSR program manager, if any: None
- Extensions granted or milestones slipped, if any: None
- Include any new discoveries, inventions, or patent disclosures during this reporting period (if none, report none): None

Robust, Adaptive Radar Detection and Estimation

Vishal Monga

Department of Electrical Engineering, The Pennsylvania State University,
University Park, PA, USA

Contents

Chapter 0

Accomplishments

Estimating the disturbance or clutter covariance is a centrally important problem in radar space time adaptive processing (STAP) since estimation of the disturbance or interference covariance matrix plays a central role on radar target detection in the presence of clutter, noise and a jammer. The disturbance covariance matrix should be inferred from training sample observations in practice. Traditional maximum likelihood (ML) estimators are effective when homogeneous (target free) training data is abundant but lead to poor estimates, degraded false alarm rates, and detection loss in the regime of limited training. However, large number of homogeneous training samples are generally not available because of difficulty of guaranteeing target free disturbance observation, practical limitations imposed by the spatio-temporal nonstationarity, and system considerations. The problem has been exacerbated by recent advances that have led to more antenna elements (J) and higher temporal resolution (P) time epochs resulting in a large dimension ($N = JP$).

In this report, we look to address the aforementioned challenges by exploiting physically inspired constraints into ML estimation. While adding constraints is beneficial to achieve satisfactory performance in the practical regime of limited training, it leads to a challenging problem. Unlike unconstrained estimators, a vast majority of constrained radar STAP estimators are iterative and expensive numerically, which prohibits practical deployment. We focus on breaking this classical trade-off between computational tractability and desirable performance measures, particularly in training starved regimes. In particular, we exploit both the structure of the disturbance covariance and importantly the knowledge of the clutter rank to yield a new rank constrained maximum likelihood (RCML) estimator of clutter/disturbance covariance. We demonstrate that the rank-constrained estimation problem can in fact be cast in the framework of a tractable convex optimization problem, and derive closed form expressions for the estimated covariance matrix. In addition, we derive a new covariance estimator for STAP that jointly considers a Toeplitz structure and a rank constraint on the clutter component. Past work has shown that in the regime of low training, even handling each constraint individually is hard and techniques often resort to slow numerically based solutions. Our proposed solution leverages the rank constrained ML estimator (RCML) of structured covariances to build a computationally friendly approximation that involves a cascade of two closed form solutions. Performance analysis using the KASSPER data set (where ground truth covariance is made available) shows that the proposed RCML estimator vastly outperforms state-of-the art alternatives even for low training including the notoriously difficult regime of $K \leq N$ training regimes and for the experiments considering real-world scenarios such as target detection performance and the case that some of training samples are corrupted by target information.

Finally, we address the problem of working with inexact physical radar parameters under a practical radar environment. As shown in this report, employing practical constraints such as a rank of the clutter subspace and a condition number of disturbance covariance leads to a practically powerful estimator as well as a closed form solution. While the rank and the condition number are very effective constraints, often practical non-ideality makes it difficult to be known precisely using physical models. We propose a robust covariance estimation method via an expected likelihood (EL) approach. We analyze covariance estimation algorithms under three different cases of imperfect constraints: 1) only rank constraint, 2) both rank and noise power constraint, and 3) condition number constraint. For each case, we formulate estimation of the constraint as an optimization problem with the expected likelihood criterion and formally derive and prove a significant analytical result such as uniqueness of the solution. Through experimental results from a simulation model and the KASSPER data set, we show the estimator with optimal constraints obtained by the EL approach outperforms alternatives in the sense of a normalized signal-to-interference and noise ratio (SINR).

Chapter 1

Rank Constrained ML Estimation of Structured Covariance Matrices

1.1 Summary

Estimating the disturbance or clutter covariance is a centrally important problem in radar space time adaptive processing (STAP). Traditional maximum likelihood (ML) estimators are effective when training data is abundant but lead to poor estimates, degraded false alarm rates, and detection loss in the realistic regime of limited training. The problem is exacerbated by recent advances which have led to high dimensionality N of the observations arising from increased antenna elements (J) as well as higher temporal resolution (P time epochs and finally $N = JP$). This work introduces physically inspired constraints into ML estimation. In particular, we exploit both the structure of the disturbance covariance and importantly the knowledge of the clutter rank to yield a new rank constrained maximum likelihood (RCML) estimator of clutter/disturbance covariance. We demonstrate that the rank-constrained estimation problem can in fact be cast in the framework of a tractable convex optimization problem, and derive closed form expressions for the estimated covariance matrix. Performance analysis using the KASSPER data set (where ground truth covariance is made available) shows that the proposed estimator vastly outperforms state-of-the art alternatives in the sense of higher normalized signal to interference and noise ratio (SINR). Crucially the proposed RCML estimator can excel even for low training including the notoriously difficult regime of $K \leq N$ training samples.

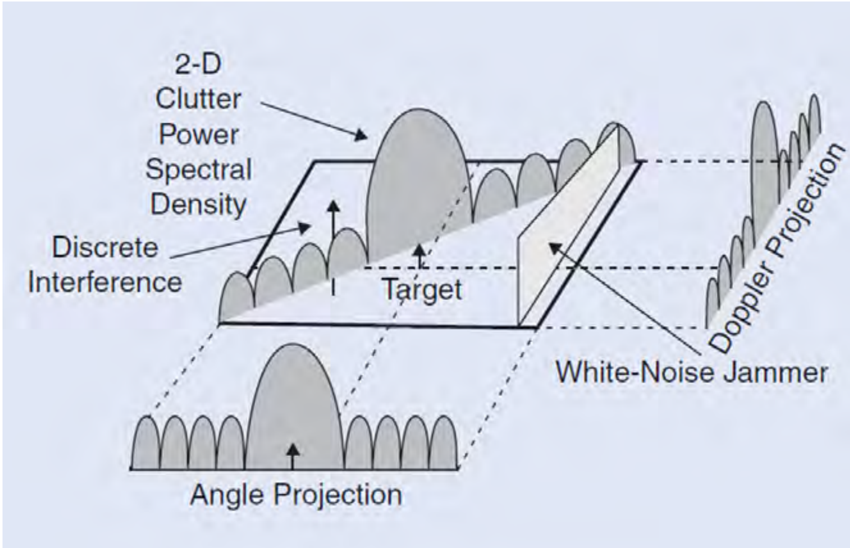


Figure 1.1: The target and interference scenario in an airborne radar.

1.2 Introduction

1.2.1 Spatio-Temporal Adaptive Processing in Radar

the radar receiver front end consists of an array of J antenna elements, which receives signals from targets, clutter, and jammers. These reflections induce a voltage at each element of the antenna array, which constitutes the measured array data at a given time instant. Snapshots of the measured data collected at P successive time epochs give rise to the spatio-temporal nature of the received radar data. The spatio-temporal product $JP = N$ is defined to be the system dimensionality. Fig. 1.1 uses the angle-Doppler space to illustrate the need for space-time adaptive processing (STAP). A target at a specific angle and traveling at a specific velocity (corresponding to a Doppler frequency) occupies a single point in this space. A jammer originates from a particular angle but is temporally (in Doppler domain) white and the clutter occupies a ridge in this 2-D space. Consequently the target signal is masked by white jammer in Doppler domain and the clutter in spatial domain and therefore, carrying out merely temporal (Doppler domain) or spatial (angle domain) processing fails to separate the target from the interference [1]. On the other hand, joint domain processing in angle and Doppler enables target detection as shown in Fig. 1.1.

The target detection problem can be cast in the framework of a statistical hypothesis test of the form

$$H_0 : \mathbf{x} \;=\; \mathbf{d} \;=\; \mathbf{c} + \mathbf{j} + \mathbf{n} \tag{1.1}$$

$$H_1 : \mathbf{x} \;=\; \alpha\mathbf{s}(\theta_t, f_t) + \mathbf{d} \;=\; \alpha\mathbf{s}(\theta_t, f_t) + \mathbf{c} + \mathbf{j} + \mathbf{n} \tag{1.2}$$

where $\mathbf{x} \in \mathbb{C}^{JP\times 1}$ is a vector form of the received data under either hypothesis, $\mathbf{d}$ represents the overall disturbance which is the sum of $\mathbf{c}$, clutter, $\mathbf{j}$, jammers, and $\mathbf{n}$, the background white noise. The vector $\mathbf{s}$ is a known spatio-temporal steering vector that represents the signal returned from the target for a specific angle and Doppler and α is the unknown target complex amplitude.

The whiten-and-match filter (MF) for detecting a rank-1 signal is the optimum processing method for Gaussian interference statistics. It is given by [2]

$$\mathbf{w} = \frac{\mathbf{R}_d^{-1}\mathbf{s}}{\sqrt{\mathbf{s}^H\mathbf{R}_d^{-1}\mathbf{s}}} \Rightarrow \Lambda_{MF} = \frac{|\mathbf{s}^H\mathbf{R}_d^{-1}\mathbf{x}|^2}{\mathbf{s}^H\mathbf{R}_d^{-1}\mathbf{s}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{MF} \tag{1.3}$$

where $\mathbf{R}_d$ is a known interference covariance matrix. Eq. (1.3) represents the matched filtering of the whitened data $\breve{\mathbf{x}} = \mathbf{R}_d^{-1/2}\mathbf{x}$ and whitened steering vector $\breve{\mathbf{s}} = \mathbf{R}_d^{-1/2}\mathbf{s}$. From Eq. (1.3), it turns out that the interference covariance matrix $\mathbf{R}_d$ is crucial in the detection statistic.

1.2.2 Motivation and Review

Because the covariance matrix plays a crucial role in the detection statistic (see Eq. (1.3)), it is very important to estimate it reliably. Widrow et al. and Applebaum proposed least-squares method [3] and maximum signal-to-noise-ratio criterion [4], respectively, using feedback loops, respectively. However, these methods were slow to converge to the steady-state solution. Reed, Mallet, and Brennan [5] verified that the sample matrix inverse (SMI) method demonstrated considerably better convergence. In the sample matrix inverse method, the disturbance covariance matrix can be estimated using K data ranges for training

$$\hat{\mathbf{R}}_d = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}_k\mathbf{x}_k^H = \frac{1}{K}\mathbf{X}\mathbf{X}^H \tag{1.4}$$

where K is the number of training data we received, $\mathbf{x}_k \in \mathbb{C}^N, N = JP$ is the kth vector of training data, and $\mathbf{X} = [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_K] \in \mathbb{C}^{N\times N}$. It is well known that the sample covariance is the *unconstrained* maximum likelihood estimator when $K \geq N$. Despite this virtue, there remain fundamental problems with the SMI approach. First, typically $K > N$ training samples are needed to guarantee the non-singularity of the estimated covariance matrix. In fact, when $K < N$ the estimate is singular and cannot be inverted which is highly undesirable in STAP. As much past research as shown [6], the estimate also does quite poorly in the vicinity of $K = N$ training samples.

large number of homogenous training samples are generally not available [7]. One reason is that it is hard to guarantee target free disturbance observations. There are also severe practical limitations imposed by the spatio-temporal nonstationarity of the interference as well as by system considerations such as bandwidth and fast scanning arrays. It is well known that $K = 2N$ training samples are needed to keep the performance within 3dB of the optimal processor. For example with $J = 11$ and $P = 32$, the parameters for the Knowledge Aided Sensor Signal Processing Expert Reasoning (KASSPER) program dataset [8], $K = 2JP = 704$ training samples are needed. Assuming an instantaneous RF bandwidth of 500 kHz, this requires wide-sense stationarity over a 400 km range. There are additional factors which make the training data scarce. They are 1.) system errors like aircraft crabbing and internal clutter motion [1], 2.) environmental considerations such as strong clutter discretes [9] and range varying interference spectra and power levels [10], and 3.) outlier contamination of training data by target-like signals [11].

To overcome the practical issue of lack of generous training, researchers have developed approaches that reduces the spatio-temporal DOF, which results in reductions in the number of required training samples and computation cost as well [6]. These works are shown in the Joint Domain Localized (JDL) processing algorithm [12], the Parametric Adaptive Matched Filter (PAMF) [13] and references therein, the Multi-Stage Wiener Filter (MSWF) [14], and factored STAP methods [1]. Another important approach is the Direct Data Domain (D^3) approach [15] which is not dependent on any statistical training. In [16], authors extend the D^3 algorithm to include statistical processing.

Finally, using signal processing and statistical learning techniques, covariance matrix estimation techniques that enforce and exploit particular structure have been pursued. Examples of structure include persymmetry [17], the Toeplitz property [18, 19, 20], circulant structure [21], multichannel autoregressive models [13, 22] and physical constraints [23]. The FML method [24] which enforces special eigenstructure also falls in this category and in fact is shown to be the most competitive technique experimentally [11, 6]. In particular, the disturbance covariance matrix $\mathbf{R}$ represents the exhibits the following structure

$$\mathbf{R} = \sigma^2\mathbf{I} + \mathbf{R}_c \tag{1.5}$$

where $\mathbf{R}_c$ denotes the clutter matrix which has a low rank and is positive semi-definite and $\mathbf{I}$ is an identity matrix. Steiner and Gerlach's FML technique ensures that the estimated covariance matrix has eigenvalues all greater than σ^2 by assuming that its value is known (or at least can be approximately known *a priori*) - which is sometimes unrealistic. Recently, the work by Aubry *et al* [25] has also improved upon FML by the introduction of a condition number constraint. Other approaches include Bayesian covariance matrix estimators [26, 27, 28, 29, 30] and the use of knowledge-based covariance models [31, 32, 33, 34]. Finally, shrinkage estimation methods have been also considered [35, 36, 37, 38].

1.3 Methods, Assumptions, and Procedures

1.3.1 Overview of Contribution

The principal contribution of our work us to incorporate the rank of the clutter component $\mathbf{R}_c$ explicitly into ML estimation of the disturbance covariance matrix. Under ideal conditions (no mutual coupling between array elements and no internal clutter motion), Brennan rule [1] states that the rank of $\mathbf{R}_c$ in the airborne linear phased array radar problem is given by

$$\text{rank}(\mathbf{R}_c) = J + \gamma(P - 1) \tag{1.6}$$

where $\gamma = 2v_pT/d$ is the slope of the clutter ridge, with v_p denoting the platform velocity, T denoting the pulse repetition interval, and d denoting the inter-element spacing. Even if there is mutual coupling in practice, $\mathbf{R}_c$ has rank r which is much less than the spatio-temporal product $N = JP$ in many practical airborne radar applications. In addition, powerful techniques have been developed [11] to determine the rank fairly accurately.

We first set up the optimization problem to estimate the disturbance covariance matrix with a structural constraint on $\mathbf{R}$ and the rank constraint on $\mathbf{R}_c$. The estimation problem when seen as an optimization over $\mathbf{R}$ is unfortunately not a convex problem, since neither the cost function nor the constraints (rank) are convex (elaborated upon in Section 1.3.2). We will however show that using a transformation of variables, reduction to a convex form is possible and further by invoking KKT conditions [39] for the resulting convex problems, it is in fact possible to derive a closed form solution. Akin to FML, we assume that the noise power σ^2 is known while setting up and solving the problem.

1.3.2 ML Estimation

Let $\mathbf{z}_i \in \mathbb{C}^N$ be the ith realization of the target-free (stochastic) disturbance vector and K be the number of training samples. That is, $i = 1, 2, \ldots, K$ and $N = JP$. Therefore, under each training sample, $\mathbf{z}_i$, under assumption of zero mean, obeys

$$f(\mathbf{z}_i) = \frac{1}{\pi^N|\mathbf{R}|}\exp(-\mathbf{z}_i^H\mathbf{R}^{-1}\mathbf{z}_i) \tag{1.7}$$

which comes from a zero-mean complex circular Gaussian distribution and $\mathbf{R}$ is the $N \times N$ disturbance covariance matrix. Further, $|\mathbf{R}|$ denotes the determinant of $\mathbf{R}$, $\mathbf{z}_i^H$ is the Hermitian (conjugate transpose) of $\mathbf{z}_i$. Let $\mathbf{Z}$ be the $M \times K$ complex matrix whose i-th column is the observed vector $\mathbf{z}_i$. Since each observations $\mathbf{z}_i$ are i.i.d, the likelihood of observing $\mathbf{Z}$ given $\mathbf{R}$ is given by

$$\begin{aligned} f_{(\mathbf{R})}(\mathbf{Z}) &= \frac{1}{\pi^{NK}}|\mathbf{R}|^{-K}\exp\big(-tr\{\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\}\big) && (1.8)\\ &= \frac{1}{\pi^{NK}}|\mathbf{R}|^{-K}\exp\big(-tr\{\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}^H\}\big) && (1.9)\\ &= \frac{1}{\pi^{NK}}|\mathbf{R}|^{-K}\exp\big(-K\cdot tr\{\mathbf{R}^{-1}\mathbf{S}\}\big) && (1.10) \end{aligned}$$

where $\mathbf{S} = \frac{1}{K}\mathbf{Z}\mathbf{Z}^H$ is the well-known sample covariance matrix. Our goal is to find the positive definite matrix $\mathbf{R}$ that maximizes the likelihood function $f_{(\mathbf{R})}(\mathbf{Z})$. The logarithm of the likelihood term is

$$\log f_{(\mathbf{R})}(\mathbf{Z}) = -K \cdot tr\{\mathbf{R}^{-1}\mathbf{S}\} - K\log(|\mathbf{R}|) - NK\log(\pi). \tag{1.11}$$

Maximizing the log-likelihood as a function of $\mathbf{R}$ is equivalent to minimizing the function given by

$$tr\{\mathbf{R}^{-1}\mathbf{S}\} + \log(|\mathbf{R}|). \tag{1.12}$$

Therefore, Eq. (1.12) is the cost function of our optimization problem. Since the cost function is not a convex function in $\mathbf{R}$, we apply a transformation variables i.e., let $\mathbf{X} = \sigma^2\mathbf{R}^{-1}$ and $\mathbf{S}' = \dfrac{1}{\sigma^2}\mathbf{S}$. Then, the revised cost function in the optimization variable $\mathbf{X}$ becomes

$$\begin{aligned} & tr\{\mathbf{R}^{-1}\mathbf{S}\} + \log(|\mathbf{R}|) \\ = \quad & tr\{\mathbf{S}'\mathbf{X}\} - \log(|\frac{1}{\sigma^2}\mathbf{X}|) \qquad (1.13) \\ = \quad & tr\{\mathbf{S}'\mathbf{X}\} - \log(|\mathbf{X}|) + \log\sigma^{2N}. \qquad (1.14) \end{aligned}$$

Since $\log\sigma^{2N}$ in Eq. (1.14) is a constant, the final cost function to be minimized is

$$tr\{\mathbf{S}'\mathbf{X}\} - \log(|\mathbf{X}|). \tag{1.15}$$

Note that $tr\{\mathbf{S}'\mathbf{X}\} = \sum_{i=1}^{N}\sum_{j=1}^{N} s'_{ji}x_{ij}$ is affine and $\log(|\mathbf{X}|)$ is concave, which implies $-\log(|\mathbf{X}|)$ is convex. Therefore, the final cost function, Eq. (1.15) is convex in the variable $\mathbf{X}$.

We now express $\mathbf{X}$ and $\mathbf{S}'$ in terms of their eigenvalue decomposition, i.e., $\mathbf{X} = \mathbf{\Phi\Lambda\Phi}^H$ and using the eigendecompositions, the cost function can be simplified as

$$\mathbf{d}^T\boldsymbol{\lambda} - \mathbf{1}^T\log\boldsymbol{\lambda} \tag{1.16}$$

where $\mathbf{d}$ and $\boldsymbol{\lambda}$ are vectors with entries of eigenvalues of $\mathbf{S}'$ and $\mathbf{X}$ respectively. This result is in fact fairly well known from standard unconstrained ML estimation of non-singular $\mathbf{R}$.

We assume the noise power is known. Then the constraints of the optimization problem are

$$\begin{cases} \mathbf{R} = \sigma^2\mathbf{I} + \mathbf{R}_c \\ rank(\mathbf{R}_c) = r \\ \mathbf{R}_c \succeq \mathbf{0} \\ \mathbf{R} \succeq \sigma^2\mathbf{I} \end{cases}. \tag{1.17}$$

Since $rank(\mathbf{R}_c) = r$, $\mathbf{R}_c$ has r non-negative eigenvalues and the rest eigenvalues are all zero. Hence, from Eq. (1.5), $\mathbf{R}$ has r eigenvalues which are greater than or equal to σ^2 and the rest eigenvalues equal to σ^2. Hence, the eigenvalues of $\mathbf{X}$ should be satisfy

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r \leq \lambda_{r+1} = \cdots = \lambda_N = 1 \tag{1.18}$$

Now the final optimization problem can be expressed in vector-matrix form,

$$\begin{cases} \min\limits_{\boldsymbol{\lambda}} & \mathbf{d}^T\boldsymbol{\lambda} - \mathbf{1}^T \log \boldsymbol{\lambda} \\ s.t. & \mathbf{F}\boldsymbol{\lambda} \preceq \mathbf{g} \\ & \mathbf{E}\boldsymbol{\lambda} = \mathbf{h} \end{cases} \tag{1.19}$$

where $\mathbf{F} = \begin{bmatrix} \mathbf{U} \\ -\mathbf{I} \\ \mathbf{I} \end{bmatrix}$, $\mathbf{g} = \begin{bmatrix} \mathbf{0} \\ -\boldsymbol{\varepsilon} \\ \mathbf{1} \end{bmatrix}$, $\mathbf{E} = \begin{bmatrix} \mathbf{0}_{r\times r} & \mathbf{0}_{r\times(N-r)} \\ \mathbf{0}_{(N-r)\times r} & \mathbf{I}_{N-r} \end{bmatrix}$, and $\mathbf{h} = [0, 0, \cdots, 0_r, 1, 1, \cdots, 1]^T$. Here, $\boldsymbol{\lambda}$, $\mathbf{d}$, $\mathbf{h} \in \mathbb{R}^N$, $\mathbf{g} \in \mathbb{R}^{3N}$, $\mathbf{U}$, $\mathbf{E} \in \mathbb{R}^{N\times N}$, and $\mathbf{F} \in \mathbb{R}^{3N\times N}$. The optimization problem (1.19) is obviously a convex optimization problem because the cost function is a convex function and feasible constraint sets are convex as well.

A closed form solution for (1.19) can in fact be derived using KKT conditions [39] in constrained optimization. The optimal solution $\boldsymbol{\lambda}^\star$ is

$$\lambda_i^\star = \begin{cases} \min(1, \dfrac{1}{d_i}) & \text{for } i = 1, 2, \ldots, r \\ 1 & \text{for } i = r+1, r+2, \ldots, N \end{cases} \tag{1.20}$$

1.4 Results and Discussions

1.4.1 Experimental Setup and Methods Compared

Data from the L-band data set of the Knowledge Aided Sensor Signal Processing and Expert Reasoning (KASSPER) program [8] is used for the performance analysis discussed in this section. The KASSPER data is the result of a significant effort by DARPA to provide a publicly available resource for the evaluation and benchmarking of radar STAP algorithms. As elaborated in [29], the KASSPER data set was carefully captured to represent real-world ground clutter and captures variations in underlying terrain, foliage and urban/manmade structures. Further, the KASSPER data set exhibits two very desirable characteristics from the viewpoint of evaluating covariance estimation techniques: 1.) the low-rank structure of clutter in KASSPER has been verified by researchers before [11, 29], and 2.) the true covariance matrices for each range bin have been made available - this facilitates comparisons via powerful figures of merit where the theoretical upper/lower bounds are known.

The L-band data set consists of a data cube of 1000 range bins corresponding to the returns

Table 1.1: KASSPER Dataset-1 parameters

Parameter	Value
Carrier Frequency	1240 MHz
Bandwidth (BW)	10 MHz
Number of Antenna Elements	11
Number of Pulses	32
Pulse Repetition Frequency	1984 Hz
1000 Range Bins	35 km to 50 km
91 Azimuth Angles	87°, 89°, ... 267°
128 Doppler Frequencies	-992 Hz, -976.38 Hz, ..., 992 Hz
Clutter Power	40 dB
Number of Targets	226 (200 detectable targets)
Range of Target Dop. Freq.	-99.2 Hz to 372 Hz

from a single coherent processing interval from $11(= J)$ channels and $32(= P)$ pulses. Therefore, the dimension of observations (or the spatio-temporal product) N is $11 \times 32 = 352$. Other key parameters are detailed in Table 2.1. Finally, a clutter rank[1] of $r = J + P - 1 = 42$ was used by our RCML estimator in all the results to follow, unless explicitly stated otherwise.

We evaluate and compare four different covariance estimation techniques:

- **Sample Covariance Matrix:** The sample covariance matrix is given in Eq. (1.4). It is well known that the sample covariance is the unconstrained maximum likelihood estimator under Gaussian disturbance statistics. Consistent with radar literature [5], we'll refer to the use of this technique as SMI.

- **Fast Maximum Likelihood:** The fast maximum likelihood (FML) [24] uses the structural constraint of the covariance matrix which is given in Eq. (1.5). The FML method just involves calculating the eigenvalue decomposition of the sample covariance and perturbing eigenvalues to conform to the structure in Eq. (1.5). The noise variance σ^2 is assumed known or pre-estimated. FML's success in radar STAP is widely known [11, 40, 6].

- **Leave-one-out shrinkage estimator:** Shrinkage estimators are powerful estimators of covariance for high dimensional data that are known to also perturb the eigenstructure of the sample covariance matrix[2] [35] - often to ensure non-singularity of the estimated covariance. While a variety of shrinkage techniques are known [35, 36, 37, 38], we choose the leave-one-out covariance matrix estimate (LOOC) shrinkage estimator [41],

$$\mathbf{R} = \beta diag(\mathbf{S}) + (1 - \beta)\mathbf{S} \tag{1.21}$$

The value of β is determined via a cross-validation technique so that the average likelihood of

[1] We set clutter ridge parameters so that $\gamma = 1$.
[2] Via this definition, the FML and RCML can also been seen as a special class of shrinkage estimators.

omitted samples is maximized. We pick this estimator because it is most suited to the problem at hand and has demonstrated success in the $K \leq N$ training regime [41].

- **Eigencanceler:** The eigencanceler (EigC) is based on the eigenanalysis which suggests a small number of eigenvalues contain all the information about interferences (jammers and clutter), and therefore, the span of the eigenvectors associated with these significant eigenvalues includes all the position vectors that comprise the interference signals [42]. Since we assume that the rank is known a priori, the eigencanceler can be compared with our estimator as we use r dominant eigenvectors as interference eigenvectors. The covariance matrix can be expressed by

$$\mathbf{R} = \sum_{i=1}^{r} p_i \mathbf{v}_i \mathbf{v}_i^H + \sigma^2 \mathbf{I} \tag{1.22}$$

  where p_i and $\mathbf{v}_i$ are the clutter power and the eigenvector corresponding to r dominant eigenvalues, respectively. For $p_i \gg \sigma^2$, it follows from [43, 11] that the estimated inverse covariance matrix can be approximated as $\hat{\mathbf{R}}^{-1} \approx \frac{1}{\sigma^2}(\mathbf{I} - \mathbf{P})$ where $\mathbf{P} = \sum_{i=1}^{r} \mathbf{v}_i \mathbf{v}_i^H$. We apply this inverse covariance matrix in computing the SINR as well as estimator variance.

- **Rank Constrained Maximum Likelihood:** Our proposed estimator (abbreviated to RCML) incorporates the structural constraint and for the first time the information of the rank of the clutter component.

1.4.2 Experimental Evaluation

The normalized signal to interference and noise ratio (SINR) is used for evaluation the aforementioned covariance estimation techniques. The SINR is desired to be as high as possible. This figure of merit is plotted against azimuthal angle as well as Doppler frequency for distinct training regimes, i.e. low, representative and generous training. We also show the plot of SINR performance versus the number of training samples. Finally, we also evaluate the robustness of our RCML estimator against perturbations in the knowledge of the true rank.

Normalized SINR vs. angle and Doppler

The normalized SINR measure [44] is commonly used in the radar literature and is given by

$$\eta = \frac{|\mathbf{s}^H \hat{\mathbf{R}}^{-1} \mathbf{s}|^2}{|\mathbf{s}^H \hat{\mathbf{R}}^{-1} \mathbf{R} \hat{\mathbf{R}}^{-1} \mathbf{s}||\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}|} \tag{1.23}$$

where $\mathbf{s}$ is the spatio-temporal steering vector, $\hat{\mathbf{R}}$ is an estimated covariance matrix, and $\mathbf{R}$ is the corresponding true covariance matrix. It is easily seen that $0 < \eta \leq 1$ and $\eta = 1$ if and only if $\hat{\mathbf{R}} = \mathbf{R}$. Since the steering vector is a function of both azimuthal angle and Doppler frequency, we evaluate the normalized SINR in both angle and Doppler domain. This would lead to a SINR surface as a function of azimuthal angle and Doppler and comparing surface plots across different covariance estimation techniques is cumbersome. We therefore obtain plots as a function of one variable (i.e. just angle/Doppler) by marginalizing (averaging) over the other variable. The SINR is plotted in dB. in all figures in this chapter, that is, $\mathrm{SINR}_{\mathrm{dB}} = 10 \log_{10} \eta$. Therefore, $\mathrm{SINR}_{\mathrm{dB}} \leq 0$.

Fig. 1.2 plots the variation of normalized SINR as a function of the azimuthal angle and the Doppler frequency for varying number of training samples, K. Specifically, Figs. 1.2 (a) and (b) are corresponding to $K = 300 < N = 352$, Figs. 1.2 (c) and (d) plots results for $K = 352 = N$, likewise Figs. 1.2 (e),(f) and (g),(h) are corresponding to $K = 750 \approx 2N$ and $K = 3000 \gg N$ respectively.

Figs. 1.2 (a),(b) and (c), (d) report results for the challenging regime of $K \leq N$. When $K < N$ the sample covariance matrix is not invertible, hence for the results in Figs. 1.2 (a),(b) we used its pseudo-inverse as a substitute. Unsurprisingly, the sample covariance technique suffers tremendously when $K \leq N$ as is evident from Figs. 1.2 (a)-(d). LOOC shrinkage does considerably better than SMI because it forces a reasonably good eigenstructure. The informed estimators, i.e. FML, EigC, and RCML perform appreciably well with RCML affording the best overall performance. It is useful to note that RCML in fact offers about 1 dB improvement over FML.

Even for representative training in Figs. 1.2 (e)-(f), the vastly superior performance of the FML, EigC, and RCML techniques is apparent. Again, by virtue of incorporating the rank information, the proposed RCML estimator outperforms the competing methods. Finally, Figs. 1.2 (g)-(h) confirm the intuition that as training becomes close to asymptotic, the gap between the various methods begins to decrease - of course, such generous training is typically impossible to obtain in practice. This is due to the fact that all the covariance matrix estimates considered converge to the true covariance matrix in the limit of large training data.

Performance vs. number of training samples

While the results in Sections 1.4.2 do explore performance against training to some extent - here we present bar graphs to explore this issue with a finer granularity. To obtain a single scalar performance measure as a function of training, averaging was carried out over both the angle and Doppler variables.

Fig. 1.3 (a) and (b) present bar graphs that quantify the SINR and estimator variance (both in dB) as a function of training samples K, where K is varied from as low as 60 to as high as 3000. Two trends are evident from Fig. 1.3 (a): 1.) as intuitively expected, the SINR values increases monotonically with an increase in the number of training samples for all methods (except for the sample covariance technique in the $K \leq N$ regime which is a well-known phenomena observed in past work as well [24]) and 2.) the RCML estimator exhibits remarkably good performance in all training regimes. Similar trends are observed for the estimator variance as well in Fig. 1.3 (b) except that we see a monotonic decrease instead. Again, the RCML estimator is consistently the best except for

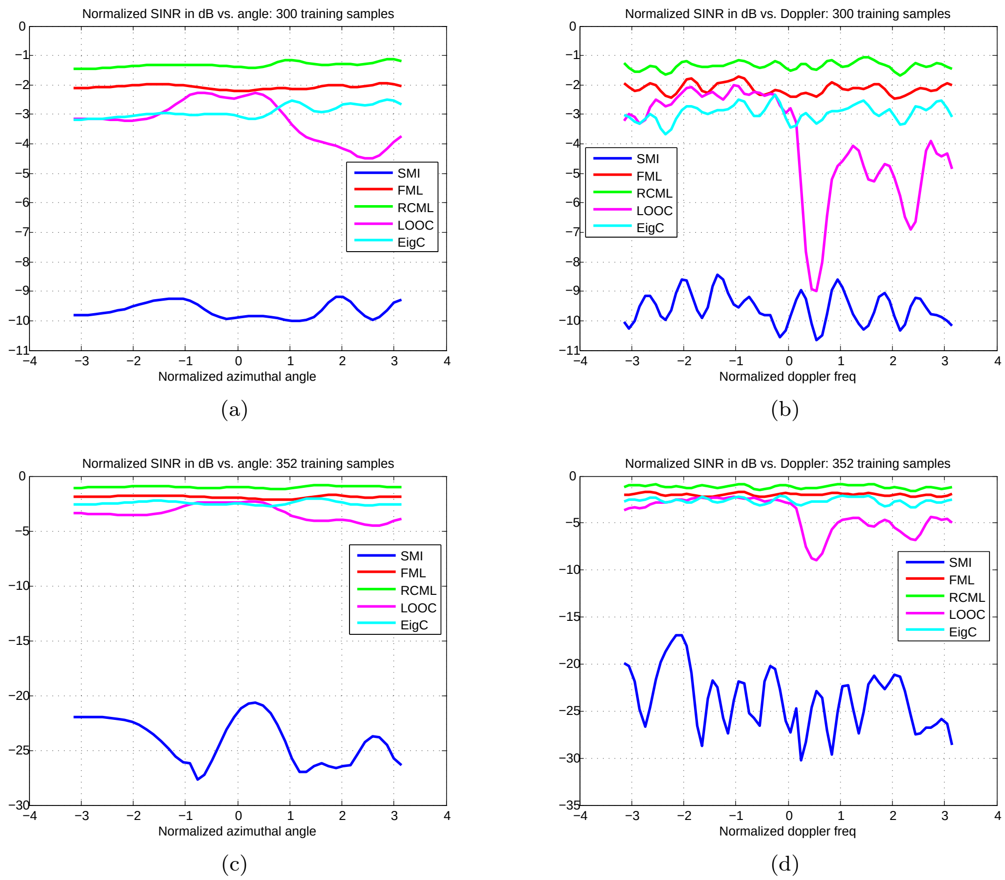


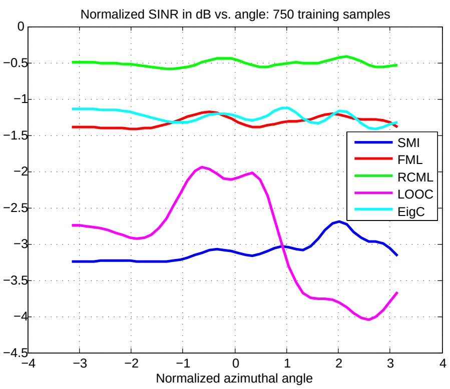
Figure 1.2: Normalized SINR vs. normalized azimuthal angle and doppler frequency, respectively (true ranges of azimuth and doppler can be seen in Table 2.1). Sample covariance matrix (SMI), fast maximum likelihood (FML), LOOC shrinkage estimator (LOOC), eigencanceler (EigC), and rank constrained maximum likelihood (RCML) estimators are the methods compared. $K = 300$ is used for (a) and (b), $K = 352$ is used for (c) and (d), $K = 750$ is used for (e) and (f), and finally $K = 3000$ is used for (g) and (h).

$K = 60$ samples where all estimators other than sample covariance are very close and no clear winner emerges.

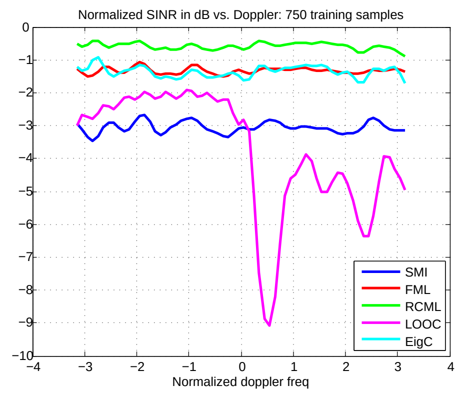
Rank Sensitivity

The KASSPER data, the clutter rank conforms to Eq. (1.6) - the Brennan rule. For the parameters used in our experiments, this would lead to a predicted ideal rank of $r = J + P - 1 = 42$. In a practical situation, departures from the ideal behavior are expected and hence we explore the performance our proposed RCML estimator even as *incorrect* rank information is used.

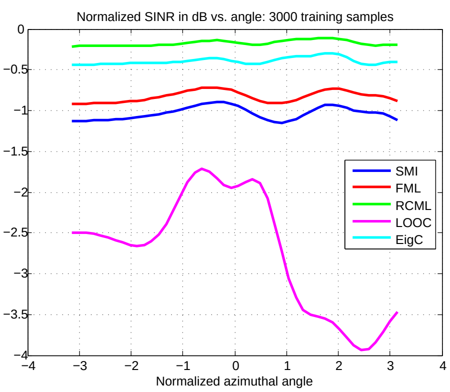
The results in Fig. 1.4 demonstrate the robustness of RCML to perturbations in the clutter rank. Fig. 1.4 presents bar graphs that show averaged SINR results for $K = 352$ and $K = 750$ training samples. We determined numerically that the "true" rank of the clutter covariance for the range



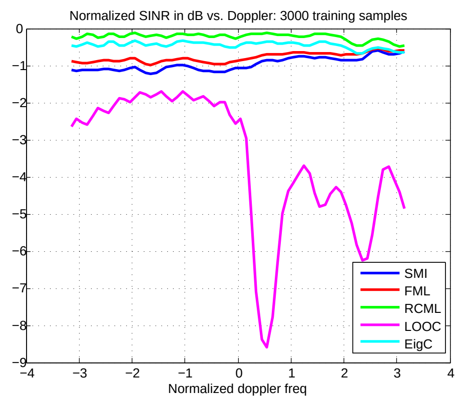
(e)



(f)



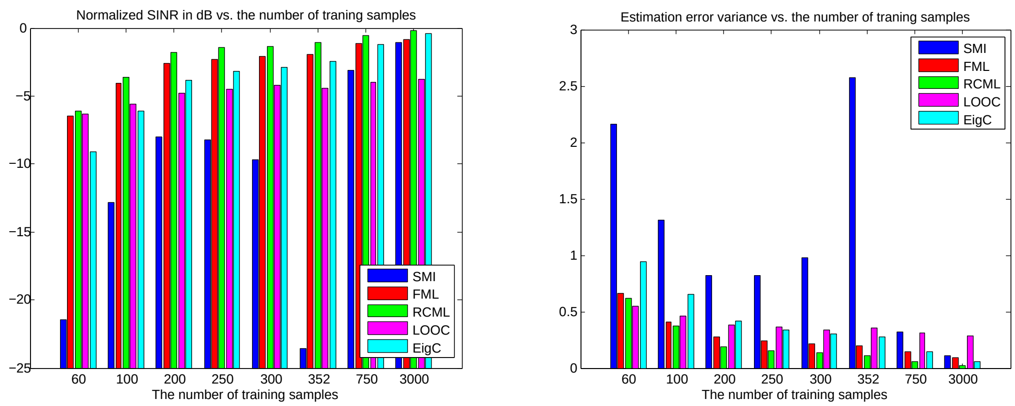
(g)



(h)

Figure 1.2: Normalized SINR vs. normalized azimuthal angle and doppler frequency, respectively (true ranges of azimuth and doppler can be seen in Table 2.1). Sample covariance matrix (SMI), fast maximum likelihood (FML), LOOC shrinkage estimator (LOOC), eigencanceler (EigC), and rank constrained maximum likelihood (RCML) estimators are the methods compared. $K = 300$ is used for (a) and (b), $K = 352$ is used for (c) and (d), $K = 750$ is used for (e) and (f), and finally $K = 3000$ is used for (g) and (h).

bin of choice was in fact 43 which is a mild departure from the 42 predicted by the Brennan rule. Comparisons are made between FML and RCML with the difference that seven variants of RCML are presented - with rank from 34 to 45. As Fig. 1.4 reveals, using the true rank of 43 indeed yields the best covariance matrix estimator but the penalty of the small departure, i.e. using a rank of 40 to 45 which are close to the true rank 43 leads to a very small performance loss. On the other hand, Figs. 1.4 also shows variants of the RCML result with a somewhat bigger departure, i.e. a rank of 34. In this case, the performance of RCML with rank 34 is appreciably lower against using rank values around the true rank 43. Remarkably, RCML with rank 34 is still competitive with FML. Overall Fig. 1.4 therefore provides two valuable insights: 1.) since rank information is predicted using the Brennan rule - small departures in practice are possible and our estimator exhibits desirable robustness against such small perturbations to rank, and 2.) the value of using the rank information is simultaneously



(a) (b)

Figure 1.3: Normalized SINR and Estimator variance vs. the number of training samples. The used numbers are 60, 100, 200, 250, 300, 352, 750, and 3000 (a) Normalized SINR and (b) Estimator error variance.

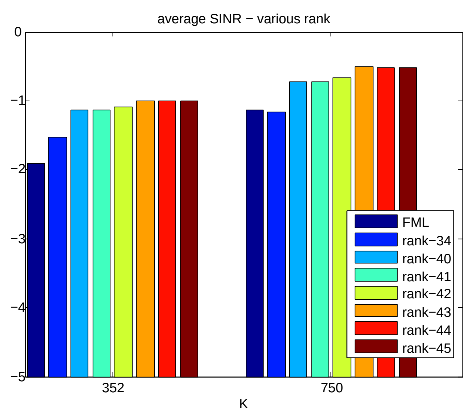


Figure 1.4: Normalized SINR of rank constrained maximum likelihood (RCML) for various rank information

revealed - because RCML with rank 34 is competitive with FML, it shows that FML significantly underestimates the true rank in these examples.

Practical Merits of the RCML Estimator

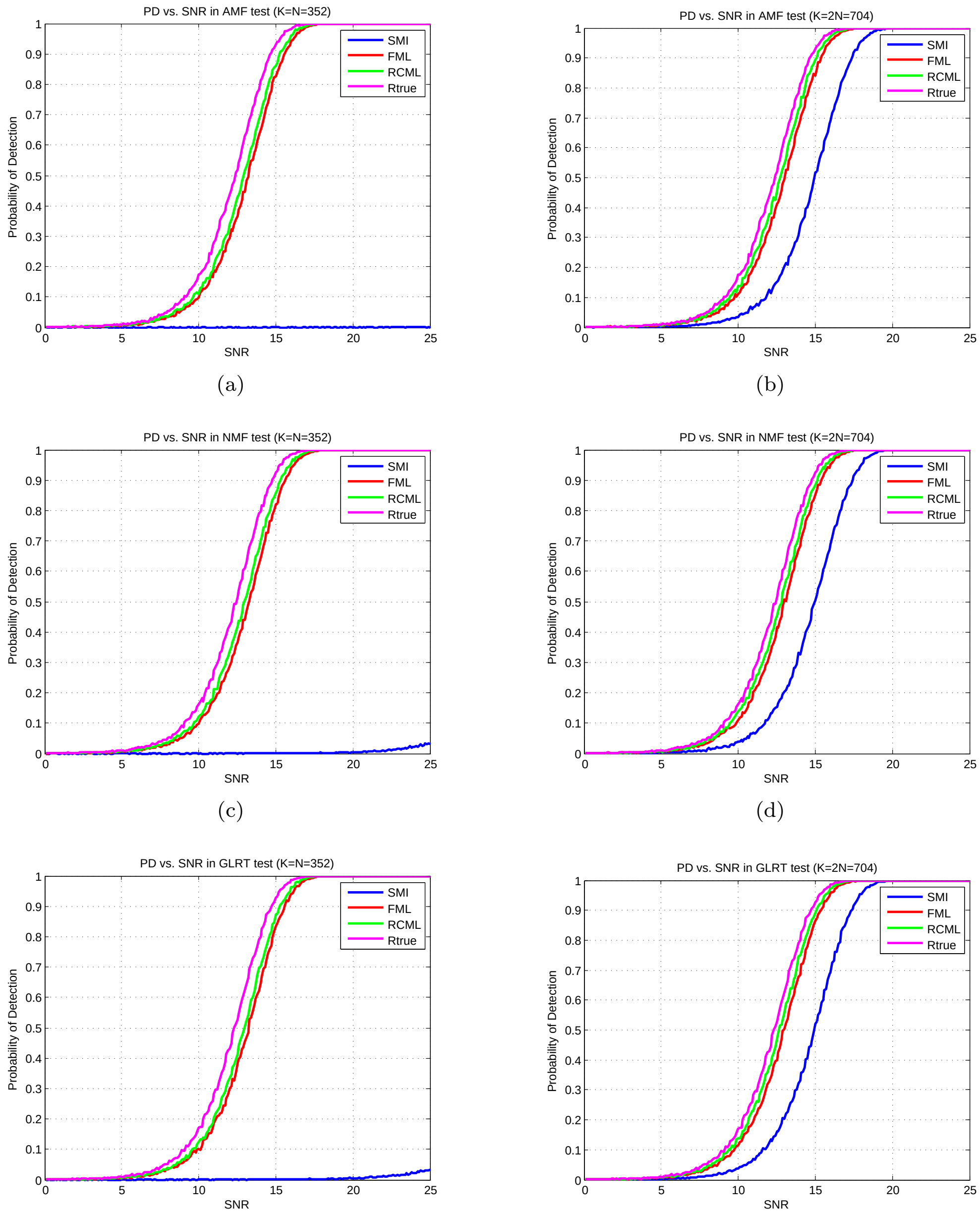
The experiments in Section 1.4.2 to Section 1.4.2, however, assume that we have access to homogeneous training samples, which is often not available in practice. This section provides more realistic and challenging practical evaluation by means of two new flavors of experimental results: 1.) plots of probability of detection versus SNR for a variety of detection statistics, and 2.) normalized SINR performance in the presence of *heterogeneous* training samples which are corrupted by the target information. Here, we consider an experimental environment which reflects real-world scenarios by considering non-homogenous training. We perform two experimental investigations. First, we examine if incorporating rank-information really leads to better target detection. Second, robustness to target contamination in training samples is investigated - while outlier removal techniques have been proposed [10, 45], in practice target contamination of training data cannot be entirely ruled out. We evaluate and compare three different covariance estimation techniques, SMI, FML and our proposed RCML. We show that the RCML estimator can still outperform alternatives in that detection probability is second only to the theoretic upper bound when the true covariance is known, and the rank information is invaluable in yielding meaningful estimates even as almost all available training samples are corrupted.

Probability of Detection vs. SNR We apply three test statistics, the normalized matched filter (NMF), the adaptive matched filter (AMF) [2], and the generalized likelihood ratio test (GLRT) [46]. The test statistics are given by

$$\begin{cases} \text{NMF:} & \dfrac{|\mathbf{s}^H\hat{\mathbf{R}}^{-1}\mathbf{e}|^2}{(\mathbf{s}^H\hat{\mathbf{R}}^{-1}\mathbf{s})(\mathbf{e}^H\hat{\mathbf{R}}^{-1}\mathbf{e})} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{\text{NMF}} \\ \text{AMF:} & \dfrac{|\mathbf{s}^H\hat{\mathbf{R}}^{-1}\mathbf{e}|^2}{\mathbf{s}^H\hat{\mathbf{R}}^{-1}\mathbf{s}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{\text{AMF}} \\ \text{GLRT:} & \dfrac{|\mathbf{s}^H\hat{\mathbf{R}}^{-1}\mathbf{e}|^2}{\mathbf{s}^H\hat{\mathbf{R}}^{-1}\mathbf{s}\left(1+\frac{1}{K}\mathbf{e}^H\hat{\mathbf{R}}^{-1}\mathbf{e}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} K\lambda_{\text{GLRT}} \end{cases} \tag{1.24}$$

where $\mathbf{s}$, $\hat{\mathbf{R}}$, $\mathbf{e}$, and K are the steering vector, the estimated covariance matrix, the observation vector, and the number of training samples, respectively. The detection probability P_d is defined as the probability that the value of test statistic is greater than a threshold conditioned on the hypothesis that the received data includes target information. Therefore, it depends on signal to noise ratio (SNR, by virtue of $\mathbf{s}$,) and the estimated covariance matrix. Since P_d does not typically admit a closed form, we first generate a number of samples from the L-band data set of KASSPER program to determine λ corresponding to the fixed false alarm rate and then employ Monte Carlo simulations to evaluate P_d corresponding to each estimator for each of the test statistics. We set a constant false alarm rate to 10^{-4}.

Fig. 1.5 shows the detection probability P_d plotted as a function of SNR for different estimators and detection statistics. Figs. 1.5 (a) and (b) plot P_d for AMF test, Figs. 1.5 (c) and (d) are corresponding to the NMF test, and Figs. 1.5 (e) and (f) plots results for the GLRT. We use $K = N = 352$ and $K = 2N = 704$ training samples to estimate the covariance matrix for each of the test statistics. Figs. 1.5 (a), (c), and (e) are for $K = 352$ and Figs. 1.5 (b), (d), and (f) are for $K = 704$. It is well-known that $K = 2N$ training samples are needed to keep the performance within 3dB. Indeed, we can see that the sample covariance matrix has about 3dB loss vs. the true covariance matrix in all of test statistics. The proposed RCML estimator is the closest to the P_d achieved by using the true covariance matrix (upper bound) and FML follows RCML. As expected, each estimator shows higher detection probability when $K = 2N$ vs. $K = N$, i.e. an increase in training. Note finally that the



(a) (b) (c) (d) (e) (f)

Figure 1.5: Probability of detection vs. SNR. Sample covariance matrix (SMI), fast maximum likelihood (FML), rank constrained maximum likelihood (RCML) estimators are the methods compared. $K = 352$ is used for (a), (c), and (e) and $K = 704$ is used for (b), (d), and (f). (a) and (b) are AMF test performance, (c) and (d) are NMF test performance, and (e) and (f) are GLRT test performance.

RCML estimator performs the best no matter which test statistic is applied and in every regime of training.

Robustness to Nonhomogeneous Training Samples We investigate two different scenarios to evaluate robustness to nonhomogeneous training samples. First, we fix the ratio of the number of

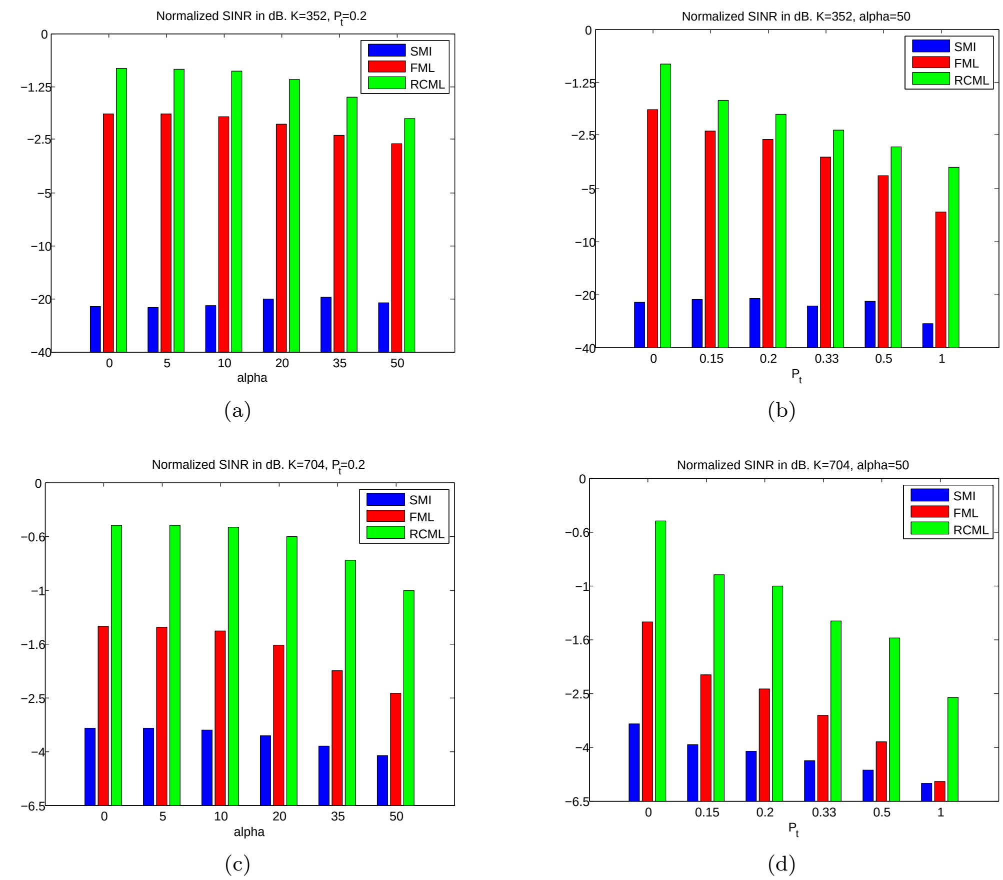


Figure 1.6: Normalized SINR in dB. vs. target intensity α and percentage corruption P_t. Sample covariance matrix (SMI), fast maximum likelihood (FML), rank constrained maximum likelihood (RCML) estimators are the methods compared. $K = 352$ is used for (a) and (b), and $K = 704$ is used for (c) and (d).

corrupted samples including target information to target-free samples. This ratio is given by

$$P_t = \frac{\text{the number of corrupted samples by target information}}{K(= \text{the number of total training samples})} \tag{1.25}$$

and the intensity of target signal by α, that is, the received data $\mathbf{z}$ can be expressed by

$$\mathbf{z} = \alpha\mathbf{s}(\theta_t, f_t) + \mathbf{d} \tag{1.26}$$

where $\mathbf{d} = \mathbf{c} + \mathbf{j} + \mathbf{n}$ represents the overall disturbance which is the sum of $\mathbf{c}$, clutter, $\mathbf{j}$, jammers, $\mathbf{n}$, the background white noise, and comes from a zero-mean complex circular Gaussian distribution. $\mathbf{s}$ is a known spatio-temporal steering vector [40] which is drawn from a distribution independent of $\mathbf{d}$. In particular, we examine performance as the percentage of corrupted samples, i.e., P_t is varied while keeping a fixed intensity of the target signal, α. Our second investigation involves varying α for a fixed P_t.

We use two evaluation measures: the normalized SINR and a trace deviation measure - TRD($\hat{\mathbf{R}}$).

1.) **Normalized SINR**: The normalized SINR measure [44] is commonly used in the radar literature and is given by

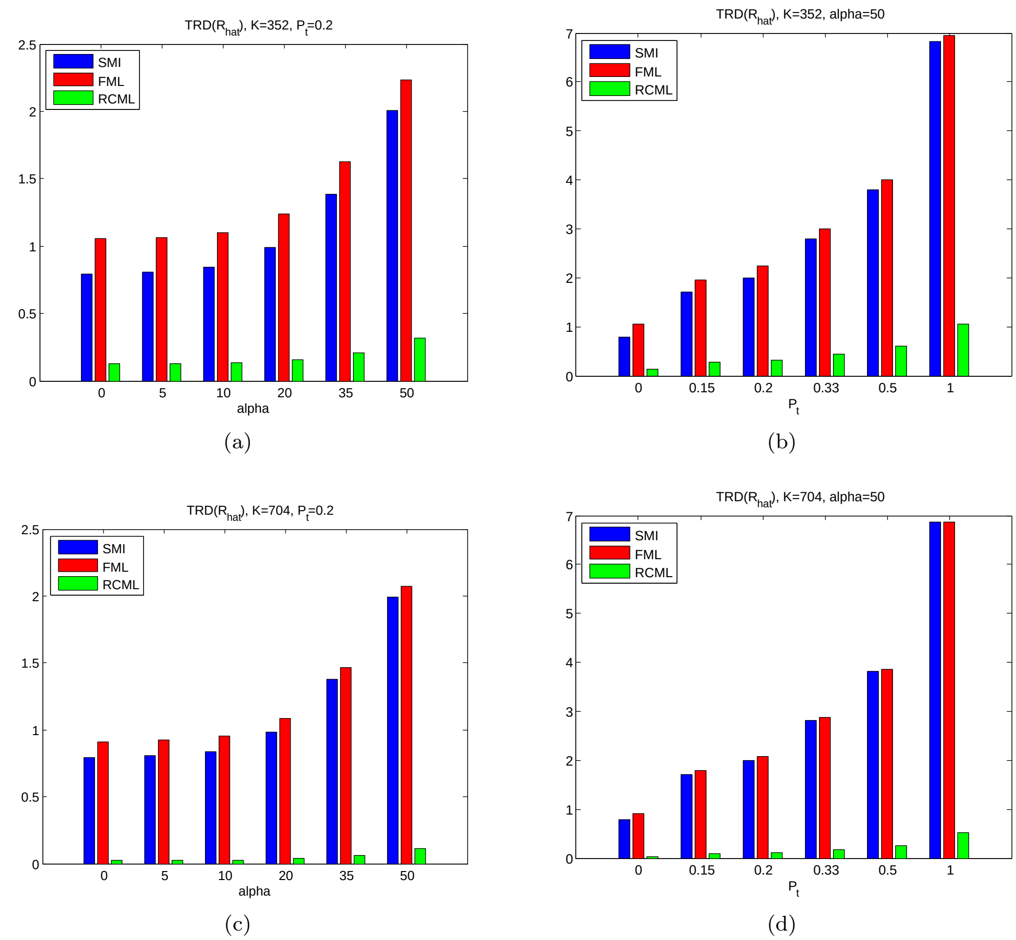
$$\eta = \frac{|\mathbf{s}^H \hat{\mathbf{R}}^{-1}\mathbf{s}|^2}{|\mathbf{s}^H \hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{s}||\mathbf{s}^H \mathbf{R}^{-1}\mathbf{s}|} \tag{1.27}$$

where $\mathbf{s}$ is the spatio-temporal steering vector, $\hat{\mathbf{R}}$ is an estimated covariance matrix, and $\mathbf{R}$ is the true covariance matrix. It is easily seen that $0 < \eta \leq 1$ and $\eta = 1$ if and only if $\hat{\mathbf{R}} = \mathbf{R}$. The SINR is plotted in dB. in all figures in this chapter, that is, $\text{SINR}_{\text{dB}} = 10\log_{10}\eta$. Therefore, $\text{SINR}_{\text{dB}} \leq 0$.

Fig. 1.6 presents bar graphs that show averaged $SINR_{dB}$ results for $K = 352$ and $K = 704$ training samples. Because the steering vector is a function of both azimuthal angle and Doppler frequency, we evaluate the normalized SINR in both angle and Doppler domain and average over both domains to get the normalized SINR value represented by each bar. Figs. 1.6 (a) and (b) are corresponding to $K = N = 352$ and Figs. 1.6 (c) and (d) plots results for $K = 2N = 704$. In particular, Figs. 1.6 (a) and (c) plot the variation of the normalized SINR for varying intensity of the steering vector α, where α is varied from as low as 0 to as high as 50. We fixed $P_t = 0.2$ in these plots. Two trends are evident from Figs. 1.6 (a) and (c): 1.) as intuitively expected, the SINR values decreases monotonically with an increase in α for all methods (except for the sample covariance technique in the $K = N$ regime) and 2.) the RCML estimator exhibits appreciably good performance in all training regimes. Figs. 1.6 (b) and (d) plot the SINR performance for varying P_t where α remains a constant, $\alpha = 50$. The range of P_t is from 0 (no target corruption) to 1 (all the samples are corrupted by target information). Similar trends are observed as well in Figs. 1.6 (b) and (d). Again, the RCML estimator consistently outperforms the other methods. An interesting observation is that $SINR_{dB}$ drops more rapidly as a function of increasing P_t vs. increasing α, which reveals that P_t is a more critical factor than α in influencing estimation with heterogeneous training.

2.) **TRD($\hat{\mathbf{R}}$)**: We define a trace deviation measure - TRD($\hat{\mathbf{R}}$) $= |tr\{\mathbf{R}\hat{\mathbf{R}}^{-1}\}/N - 1|$ that is an alternate way of evaluating the performance of covariance matrix estimators. Intuitively, we can see $tr\{\mathbf{R}\hat{\mathbf{R}}^{-1}\}/N = 1$ when $\hat{\mathbf{R}} = \mathbf{R}$. Therefore, we can say the goal of estimation is to tray and keep TRD($\hat{\mathbf{R}}$) as small as possible, ideally close to 0. Figs. 1.7 shows plots bar graphs in the same training regime as Figs. 1.6. We plots values of TRD($\hat{\mathbf{R}}$) for varying α and P_t and the number of training samples are $K = 352$ and $K = 704$.

We can observe trends similar to those in Fig. 1.6. The TRD values monotonically increase as α and P_t increase for all methods. The proposed RCML estimator consistently outperforms other techniques considered in all experiments. Additionally, the merits of RCML in robust estimation are brought out. The TRD values corresponding to both the sample covariance matrix and the FML estimator increase quite dramatically with an increase in α and, especially, P_t. However, in the case of the RCML estimator this increase is more gradual. The TRD values corresponding to RCML in



(a) (b) (c) (d)

Figure 1.7: TRD($\hat{\mathbf{R}}$) vs. target intensity α and percentage corruption P_t. Sample covariance matrix (SMI), fast maximum likelihood (FML), rank constrained maximum likelihood (RCML) estimators are the methods compared. $K = 352$ is used for (a) and (b), and $K = 704$ is used for (c) and (d).

Fig. 1.7 (d) are in fact still close to 0 even under severe target corruption, i.e. $P_t = 1$.

1.5 Conclusions

We developed a new estimator of structured covariance matrices (identity plus a positive semi-definite component) which employs rank of the positive semi-definite matrix as an explicit constraint in ML estimation. In radar applications, the rank-deficient component corresponds to the clutter and its rank can be determined using the Brennan rule for airborne radar interacting with land clutter. We demonstrated that despite the presence of the challenging rank-constraint, the estimation problem can in fact be reduced to a convex optimization problem and admits a closed form solution. Experimentally, rank information plays a vital role and rigorous evaluation over the KASSPER data set establishes merits of the proposed estimator when evaluated via powerful figures of merit such as normalized SINR and estimator variance. Further, significant practical merits are revealed in challenging real-world radar detection and estimation set-ups. Future work could consider the incorporation of more constraints on the clutter/disturbance matrix such as Toeplitz structure as well as the use of physically inspired probabilistic priors in a Bayesian setting.

Chapter 2

Efficient Approximation of Structured Covariance under Joint Toeplitz and Rank Constraints

2.1 Summary

Disturbance covariance estimation is a centrally important problem in radar space time adaptive processing (STAP). Because training is invariably scarce, estimators that exploit inherent structure and physical radar constraints are needed in practice. This chapter develops a new computationally efficient estimator which jointly enforces a Toeplitz structure and a rank constraint on the structured interference. Previous work has shown that exact ML estimation of Toeplitz covariance matrix has no closed form solution and most versions of this problem result in iterative estimators which are computationally expensive. Our proposed solution focuses on a computationally efficient approximation and involves a cascade of two closed form solutions. First, we obtain the rank constrained ML estimator (RCML) whose merits have recently been established firmly for radar STAP. The central contribution of this chapter is the rank preserving Toeplitz approximation, which we demonstrate can be modeled as an equality constrained quadratic program and also admits a closed form. Extensive performance evaluation on both simulated and KASSPER data confirms that the proposed estimator yields unbeatable performance for radar STAP under the previously stated conditions of rank and Toeplitz constraints.

2.2 Introduction

Radar systems using multiple antenna elements that coherently process multiple pulses offer significant benefits in many applications. The directivity and resolution limits of a single sensor can be overcome by using an adaptive array of spatially distributed sensors makes multiple temporal snapshots processing possible. Specifically, joint adaptive processing in the spatial and temporal domains [44, 47, 48] called space time adaptive processing (STAP) creates an ability to suppress interference signals while simultaneously preserving gain on the desired signal. For STAP to be successful though, interference statistics, in particular the covariance matrix of the disturbance or interference must be estimated from target free training data, and therefore training plays a pivotal role in adaptive radar systems.

To obtain accurate estimates of the disturbance covariance matrix, a large number of homogeneous (target free) disturbance training samples are required in the absence of any prior knowledge about the interference environment. A compelling challenge for radar STAP emerges since generous homogeneous training is often not available in practice [7]. This problem is exacerbated because the estimation process must be repeated for each range bin of interest. Much recent research in radar STAP has been proposed to overcome the lack of generous homogeneous training. One approach to this problem uses *a priori* information about the radar environment and is widely referred to in the literature as knowledge-based processing [29, 49, 50, 40, 51, 33, 34, 52]. A subset of this technique deals with intelligent training selection for reducing both the number of required training samples and computational cost [40, 12, 6]. Another approach to improve the target detection performance is data selection screening among the training data to excise potential outliers [53, 54].

Covariance matrix estimation techniques that enforce and exploit specific structure inherent to the disturbance phenomenon have merit in the regime of extremely limited training data. Examples of structure include persymmetry [17], eigenstructure [24, 42], circulant structure [21], rank constraint [55, 56], multichannel autoregressive models [13, 22], physical constraints [23] and so on. In particular, since the covariance matrix from a stationary stochastic signal is Hermitian and Toeplitz, estimating Toeplitz covariance benefits many applications such as array processing and time series analysis. Such a Hermitian Toeplitz matrix models the covariance of a random vector obtained by sampling a wide sense stationary noise field with a uniform linear array and uncorrelated narrow-band interferers [18]. The seminal work by Burg *et al.* [57] proposed an *iterative* method for estimation of structured covariance matrices using the ML method in its full generality . Li *et al.* developed the asymptotic maximum likelihood (AML) estimation for structured covariance matrices [19] using the extended invariance principle (EXIP) [58]. Approximation of arbitrary matrices by a (Hermitian) Toeplitz matrix using matrix decompositions and outer approximations has separately been pursued in applied mathematics [59, 60, 61, 62]. While the techniques in [59, 60, 61, 62] were not conceived for signal processing or radar STAP, they can potentially be used in conjunction with classical covariance estimation. Of particular interest is Al-Homidan's l_1 sequential quadratic programming (SQP) method to find the nearest symmetric positive semi-definite Toeplitz matrix to given a matrix [59].

2.2.1 Motivation and Challenges

Various estimation and approximation techniques of Toeplitz covariance matrices have been proposed [63, 64, 65, 66]. It is well known [18] though that there is no closed-form solution for the ML estimation of a Hermitian Toeplitz covariance matrix. Many Toeplitz covariance estimation techniques need the assumption of large sample size (i.e. observed training) for computational tractability [19],[66]. In the regime of realistic training, methods rely on numerical optimization (often non-convex), are computationally involved and hence unsuitable for real-time/practical deployment.

Previous works, notably in statistics [67, 68] (and references therein) have also shown that the rank of the structured interference can be exploited in a tractable manner. Rank is a powerful constraint in covariance estimation and can often be determined via underlying radar physics. Under nominal assumptions, the Brennan rule [1] may be used to determine the rank of the structured interference. Related work also addresses the problem of determining rank in non-ideal scenarios [69]. Recently, Kang *et al.* proposed the rank constrained ML (RCML) estimation of structured covariance matrices [70] which exploits the knowledge of the radar noise floor. Kang *et al.* [70] also report another estimator called $\mathrm{RCML}_{\mathrm{LB}}$ for the case when the noise floor is assumed unknown and only a lower bound (LB) is available. The $\mathrm{RCML}_{\mathrm{LB}}$ estimator generalizes the well-known result in statistics [67, 68]. In the radar context though, the noise variance is assumed known since it can be determined by placing the radar in receive only mode [71]. Notable contributions which deal with both the rank information and Toeplitz structure of the covariance matrix jointly includes the iterated Toeplitz approximation method (ITAM) [72] proposed by Wilkes and Hayes and the iterative approach by Forster *et al.* [65]. Both approaches are based on a computationally expensive iterative procedure. The ITAM estimator in particular has been shown to be effective under very low training because of its ability to exploit structure but does not yield scalable performance improvements as realistic or generous training is made available.

2.2.2 Our contributions

It may be inferred that for adequate performance under limited training, computationally involved estimators such as ITAM [72] are needed but online covariance estimation is often needed in near real-time. While fast, closed form estimators such as AML [19] can be used, they do not excel under low or realistic training. Our contribution aims to break this classical trade-off. We develop a computationally efficient approximation of structured covariance under joint Toeplitz and rank (EASTR) constraints. Specifically, our key contributions are listed next.

- **Analytically tractable framework for exploiting both Toeplitz structure and the rank of the structured interference.** Our proposed estimator, i.e. EASTR, satisfies both Toeplitz structure property (at least approximately) and the rank information of the structured interference at the same time. Decades of research has shown that enforcing even each constraint individually can be quite onerous (e.g. rank is a non-convex constraint and no known closed form exists under the Toeplitz constraint for all training). The rank constrained ML estimation has been achieved recently though [70] via a transformation of variables. However, this does not apply when the Toeplitz constraint is added. We propose to decouple the rank and Toeplitz constraints, which lends analytical tractability. Crucially, the EASTR solution does not need iterative steps like ITAM and as will be established in Section 2.4, Furthermore, our results demonstrate that EASTR consistently outperforms ITAM.

- **Computationally efficient and fast estimation and approximation.** Our proposed method, EASTR, essentially involves a cascade of two steps where a closed form solution is available in each step. First a closed form solution using maximum likelihood employing the rank constraint is obtained from the RCML [70] estimator. Next, we propose a new method to perturb the eigenvalues of the RCML estimator in a rank preserving manner so as to impose the Toeplitz structure. We formulate a new quadratic programming (QP) optimization problem that solves for the eigenvalues while incorporating Toeplitz constraints and demonstrate that this problem also admits a closed form solution.

- **Experimental insights and improved performance in low training regimes.** The merits of EASTR are also verified experimentally over both simulated data and realistic data sets such as Knowledge Aided Sensor Signal Processing and Expert Reasoning (KASSPER). ITAM works well particularly in low training regimes but is numerically expensive. The asymptotic ML estimation gives us a fast closed form solution but shows good performances only in high training regimes. EASTR excels across all training regimes while still permitting closed form solutions attractive for practical deployment.

We consider two cases: 1.) when the Toeplitz constraint is satisfied exactly, we obtain the exact Toeplitz estimate satisfying the rank constraint and Toeplitz property and 2.) when the Toeplitz constraint is not exactly satisfied, we make slight modification on the Toeplitz constraint and derive a modified optimization problem to obtain approximately Toeplitz estimate. In practice, the available data dictates which of the two cases in invoked. Experimental investigation shows that the EASTR can outperform alternatives in the sense of 1.) normalized SINR and 2.) the probability of detection, and 3.) a newly proposed trace deviation measure.

2.3 Methods, Assumptions, and Procedures

The maximum likelihood covariance estimate $\mathbf{R}$ is one which maximizes the likelihood function based on a zero-mean complex circular Gaussian distribution:

$$f_{(\mathbf{R})}(\mathbf{Z}) = \frac{1}{\pi^{NK}}|\mathbf{R}|^{-K}\exp\big(-tr\{\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\}\big) \tag{2.1}$$

under both Toeplitz and rank constraints. In (2.1), K is the number of training samples, N is the dimension of observations, and $\mathbf{Z}$ is an $N \times K$ matrix whose each column is an i.i.d. observation vector. With some algebraic manipulations, the final optimization problem may be written as

$$\begin{cases} \min\limits_{\mathbf{R}} & tr\{\mathbf{R}^{-1}\mathbf{S}\} + \log(|\mathbf{R}|) \\ s.t. & \mathbf{R} = \sigma^2\mathbf{I} + \mathbf{R}_c \\ & \text{rank}(\mathbf{R}_c) = r \\ & \mathbf{R}_c \in T \end{cases} \tag{2.2}$$

where $\mathbf{S} = \frac{1}{K}\mathbf{Z}\mathbf{Z}^H$ is the sample covariance matrix, $\mathbf{R}_c$ denotes the interference covariance matrix, $\mathbf{I}$ is an $N \times N$ identity matrix, and σ^2 is the radar noise floor which can be readily determined using standard techniques [71], and lastly T is the set of all $N \times N$ Hermitian positive semi-definite Toeplitz matrices,

$$T = \{\mathbf{T} : \mathbf{T} \in \mathbb{C}^{N\times N}, \mathbf{T}^H = \mathbf{T}, \mathbf{T} \succeq \mathbf{0} \text{ and } \mathbf{T} \in \mathcal{T}\} \tag{2.3}$$

where $\mathcal{T}$ is the set of all Toeplitz matrices. The optimization problem (2.2) is particularly hard to solve because 1.) the problem in not convex (and no known transformations exist to turn it into one); hence a global minimizer is virtually impossible to find, 2.) from a numerical standpoint, solutions are known to be computationally burdensome under the Toeplitz constraint alone [57], [72]. Adding the rank constraint only exacerbates the problem.

In view of the aforementioned challenges, we focus on covariance matrix estimation that: 1.) is fast and based on analytical closed forms so as to facilitate practical deployment, and 2.) exploits previously known insights in radar STAP so that performance in the sense of high SINR and P_d can be obtained across *all* training regimes.

Our proposed solution decouples the rank and Toeplitz constraints, and develops a cascade of two closed forms as the final estimator. The first closed form is obtained by employing the recently proposed rank constrained ML (RCML) estimator of structured covariance [70]. The final RCML solution is given by [70]

$$\mathbf{R}^{\star} = \sigma^2\mathbf{X}^{\star -1} = \sigma^2\mathbf{\Phi}\mathbf{\Lambda}^{\star -1}\mathbf{\Phi}^H \tag{2.4}$$

where $\mathbf{\Phi}$ is the eigenvector matrix of the sample covariance matrix $\mathbf{S}$ and $\mathbf{\Lambda}^{\star}$ is a diagonal matrix with optimal diagonal entries $\lambda_i^\star$ which is given by

$$\lambda_i^\star = \begin{cases} \min(1, \dfrac{1}{d_i}) & \text{for } i = 1, 2, \ldots, r \\ 1 & \text{for } i = r+1, r+2, \ldots, N \end{cases} \tag{2.5}$$

where d_i is the ith eigenvalue of the sample covariance matrix normalized by σ^2, $\mathbf{S}' = \frac{1}{\sigma^2}\mathbf{S}$.

2.3.1 Conditions for Eigenvalues of Toepiltz Covariance

Our approach now involves enforcing the Toeplitz structure on top of the RCML estimator in (2.5). Let the eigenvector matrix of $\mathbf{S}$ be $\boldsymbol{\Phi}$ and the eigenvalues of $\mathbf{R}_c$ be $\lambda_1, \lambda_2, \ldots, \lambda_r, \ldots, \lambda_N$. Since we want to preserve the clutter rank constraint $rank(\mathbf{R}_c) = r$, $\mathbf{R}_c$ should have only r positive eigenvalues and the rest of them should be zero, that is

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_N = 0 \tag{2.6}$$

Therefore, $\mathbf{R}_c$ can be expressed as

$$\mathbf{R}_c = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^H \tag{2.7}$$

where

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \lambda_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{2.8}$$

and

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} & \cdots & \phi_{1N} \\ \phi_{21} & \phi_{22} & \phi_{23} & \cdots & \phi_{2N} \\ \phi_{31} & \phi_{32} & \phi_{33} & \cdots & \phi_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_{N1} & \phi_{N2} & \phi_{N3} & \cdots & \phi_{NN} \end{bmatrix} \tag{2.9}$$

Therefore, we know that ijth component of $\mathbf{R}_c$ is given by

$$(\mathbf{R}_c)_{ij} = \sum_{k=1}^{r} \lambda_k \phi_{ik} \phi_{jk}^* \tag{2.10}$$

Note that $\mathbf{R}_c$ is already Hermitian, that is, $(\mathbf{R}_c)_{ij} = (\mathbf{R}_c)^*_{ji}$. Now in order for $\mathbf{R}_c$ to be Toeplitz matrix, all entries on each diagonal in the lower triangular part in $\mathbf{R}_c$ must have same values, i.e., the following equations must hold.

$$\begin{cases} (\mathbf{R}_c)_{11} = (\mathbf{R}_c)_{22} = \cdots = (\mathbf{R}_c)_{NN} \\ (\mathbf{R}_c)_{21} = (\mathbf{R}_c)_{32} = \cdots = (\mathbf{R}_c)_{N,N-1} \\ (\mathbf{R}_c)_{31} = (\mathbf{R}_c)_{42} = \cdots = (\mathbf{R}_c)_{N,N-2} \\ \vdots \\ (\mathbf{R}_c)_{N-1,1} = (\mathbf{R}_c)_{N2} \end{cases} \tag{2.11}$$

Let us examine the first condition in (2.11), $(\mathbf{R}_c)_{11} = (\mathbf{R}_c)_{22}$,

$$\sum_{k=1}^{r} \lambda_k \phi_{1k} \phi^*_{1k} = \sum_{k=1}^{r} \lambda_k \phi_{2k} \phi^*_{2k} \tag{2.12}$$

It can be also expressed as

$$\sum_{k=1}^{r} \lambda_k (\phi_{1k} \phi^*_{1k} - \phi_{2k} \phi^*_{2k}) = 0 \tag{2.13}$$

In vector form, the first equation is given by

$$\begin{bmatrix} \phi_{11}\phi^*_{11} - \phi_{21}\phi^*_{21} & \cdots & \phi_{1r}\phi^*_{1r} - \phi_{2r}\phi^*_{2r} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{bmatrix} = 0 \tag{2.14}$$

Since the elements ϕ_{ij} of the eigenvalue matrix $\mathbf{\Phi}$ are known ($\mathbf{\Phi}$ is the eigenvector matrix of the sample covariance matrix), we now have the first constraint for Toeplitz covariance matrix as a linear combination of the eigenvalues. Other equations in Eqs. (2.11) also can be expressed in a vector form as in (2.14). Consequentially, we have a total of $N(N-1)/2$ equations and finally get the following equation which is the equality constraint of our optimization problem.

$$\mathbf{\Psi} \boldsymbol{\lambda} = \mathbf{0} \tag{2.15}$$

where each row of $\mathbf{\Psi} \in \mathbb{C}^{N(N-1)/2 \times r}$ denotes coefficients of λ_i which come from each of equations in Eqs. (2.11) and $\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_r \end{bmatrix}^T$.

Since $\mathbf{\Psi}$ Eq. (2.15) is a tall matrix, (2.15) in general is a overdetermined linear system, that is, we have more equations than unknowns. The solution set therefore depends on the rank of $\mathbf{\Psi}$. The first case is that we have an infinite set of solutions when the column rank of $\mathbf{\Psi}$ is less than r. On the other hand, when $\mathbf{\Psi}$ has a full column rank, we have the trivial solution, $\boldsymbol{\lambda} = \mathbf{0}$. That is, the covariance matrix can only be made approximately (and not exactly) Toeplitz in this case - this remedy is discussed in Section 2.3.3.

2.3.2 Exact Toeplitz Solution

When the column rank of $\mathbf{\Psi}$ is less than r, Eq. (2.15) has an infinite number of solutions. In this case, we can obtain the exact Toeplitz solution. First, let $\boldsymbol{\lambda}_{\mathrm{RCML}}$ be the eigenvalues obtained from the RCML estimation, which is given by Eq. (2.5). We already know the eigenvalues $\boldsymbol{\lambda}$ from the RCML estimate are the optimal ML estimate of the true structured covariance matrix under only the rank constraint. Therefore, we want the eigenvalues of the interference covariance matrix to satisfy Eq. (2.15) and to be as close to the RCML solution as possible. Since Eq. (2.15) has an infinite number of solutions, we can find the closest vector of the eigenvalues to $\boldsymbol{\lambda}_{\mathrm{RCML}}$ by solving the following convex optimization problem.

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \quad & ||\boldsymbol{\lambda}_{\mathrm{RCML}} - \boldsymbol{\lambda}||^2 \\ \text{subject to :} \quad & \mathbf{\Psi}\boldsymbol{\lambda} = \mathbf{0} \end{aligned} \tag{2.16}$$

The optimization problem (2.16) is a well known quadratic programming (QP) optimization problem with an equality constraint and therefore the closed form solution is available using KKT condition [39] and it is given by solving the following equation.

$$\begin{bmatrix} 2\mathbf{I} & \mathbf{\Psi}^T \\ \mathbf{\Psi} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}^\star \\ \boldsymbol{\nu}^\star \end{bmatrix} = \begin{bmatrix} 2\boldsymbol{\lambda}_{\mathrm{RCML}} \\ \mathbf{0} \end{bmatrix} \tag{2.17}$$

where $\boldsymbol{\nu}^\star$ is a vector of Lagrange multipliers.

However, the matrix on the left-hand side of Eq. (2.17) is actually singular because $\mathbf{\Psi}$ has not full column rank. So we introduce a new matrix $\check{\mathbf{\Psi}}$ instead of $\mathbf{\Psi}$ to make the left matrix invertible when we solve it. That is,

$$\begin{bmatrix} 2\mathbf{I} & \check{\mathbf{\Psi}}^T \\ \check{\mathbf{\Psi}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}^\star \\ \boldsymbol{\nu}^\star \end{bmatrix} = \begin{bmatrix} 2\boldsymbol{\lambda}_{\mathrm{RCML}} \\ \mathbf{0} \end{bmatrix} \tag{2.18}$$

where $\check{\mathbf{\Psi}}$ is a matrix consists of rank($\mathbf{\Psi}$) linearly independent rows of $\mathbf{\Psi}$. Obviously, Eq. (2.17) and Eq. (2.18) have the same solution because linearly independent rank($\mathbf{\Psi}$) rows of $\mathbf{\Psi}$ determine the set of solutions of the equation and removing redundant rows does not make any changes to the solution. It follows that the final closed form solution using blockwise inversion property is given by

$$\boldsymbol{\lambda}^\star = \big(\mathbf{I} - \check{\mathbf{\Psi}}^T(\check{\mathbf{\Psi}}\check{\mathbf{\Psi}}^T)^{-1}\check{\mathbf{\Psi}}\big)\boldsymbol{\lambda}_{\mathrm{RCML}} \tag{2.19}$$

and the final covariance matrix can be obtained by

$$\mathbf{R}^\star = \sigma^2\mathbf{I} + \mathbf{\Phi} diag(\boldsymbol{\lambda}^\star)\mathbf{\Phi}^H \tag{2.20}$$

2.3.3 Toeplitz Approximation

In the case that $\mathbf{\Psi}$ has a full column rank, Eq. (2.15) has the only one solution, $\boldsymbol{\lambda} = \mathbf{0}$, which does not yield a meaningful covariance matrix. In this case, the optimization problem to enforce the Toeplitz structure must be modified. One possibility is to explicitly incorporate the eigenvector matrix into the optimization. This however, will lead to a computationally expensive problem because the optimization must constrain the eigenvector matrix to be unitary. Further, using an eigenvector matrix to agree with $\mathbf{\Phi}$, i.e. the one obtained from sample covariance has been known to be very successful in radar STAP [47, 25, 70].

We therefore take the approach of building an *approximately* as opposed to exactly Toeplitz matrix. This can be done by computing the closest rank deficient matrix $\tilde{\mathbf{\Psi}}$ to $\mathbf{\Psi}$. Consider the singular value decomposition of $\mathbf{\Psi}$,

$$\mathbf{\Psi} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H \tag{2.21}$$

The well-known theorem, Eckart-Young theorem [73], says that a matrix $\bar{\mathbf{\Psi}}$ with the column rank less than r that minimizes $||\mathbf{\Psi} - \tilde{\mathbf{\Psi}}||_F$ is given by

$$\tilde{\mathbf{\Psi}} = \mathbf{U}\tilde{\mathbf{\Sigma}}\mathbf{V}^H \tag{2.22}$$

where $\tilde{\mathbf{\Sigma}}$ is the diagonal matrix obtained from $\mathbf{\Sigma}$ by replacing the r-th diagonal element which is the smallest diagonal element by zero. By substituting $\mathbf{\Psi}$ with $\tilde{\mathbf{\Psi}}$ in Eq. (2.15), we obtain the infinite number of solutions for $\boldsymbol{\lambda}$. Now, the optimization problem becomes

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \quad & ||\boldsymbol{\lambda}_{\mathrm{RCML}} - \boldsymbol{\lambda}||^2 \\ \text{subject to :} \quad & \tilde{\mathbf{\Psi}}\boldsymbol{\lambda} = \mathbf{0} \end{aligned} \tag{2.23}$$

Finally, a Toeplitz matrix is obtained by solving the above optimization problem in the same way done in the case of exact Toeplitz solution, that is,

$$\boldsymbol{\lambda}^{\star} = \left(\mathbf{I} - \breve{\mathbf{\Psi}}^T(\breve{\mathbf{\Psi}}\breve{\mathbf{\Psi}}^T)^{-1}\breve{\mathbf{\Psi}}\right)\boldsymbol{\lambda}_{\mathrm{RCML}} \tag{2.24}$$

where $\breve{\mathbf{\Psi}}$ is a matrix consists of $r-1$ linearly independent rows of $\tilde{\mathbf{\Psi}}$.

Remark: It should be noted that the actual rank of $\mathbf{\Psi}$ which is derived from $\mathbf{\Phi}$ depends on the training data. If the true covariance is indeed Toeplitz, we expect training samples to reflect that particularly in the regime of $K >> N$ training samples (asymptotic regime), this is indeed what we observe in practice.

2.4 Results and Discussions

2.4.1 Experimental Setup and Methods Compared

In this section, we compare the performance of proposed estimator against state of the art Toeplitz STAP estimators. Two data sets are used: 1.) A radar covariance simulation model and 2.) the well known KASSPER [8] data set.

First, we model a radar system with an N-element uniform linear array. The overall disturbance is composed of jammer and white interference. Therefore, the external wideband noise environment via its input covariance matrix can been modeled by

$$\mathbf{R}(n,m) = \sum_{i=1}^{J} \sigma_i^2 \operatorname{sinc}[0.5\beta_i(n-m)\phi_i] e^{j(n-m)\phi_i} + \sigma_a^2 \delta(n,m) \tag{2.25}$$

where $n, m \in \{1, \ldots, N\}$, J is the number of jammers, σ_i^2 is the power associated with the ith jammer, ϕ_i is the jammer phase angle with respect to the antenna phase center, β_i is the fractional bandwidth, σ_a^2 is the actual power level of the white disturbance term, and $\delta(n,m)$ has the value of 1 only when $n = m$ and 0 otherwise. This simulation model has infact been widely and very successfully used in previous literature [24, 25, 74, 33] for performance analysis. It is easily seen that $\mathbf{R}$ is Hermitian and Toeplitz since $\mathbf{R}(n,m)$ depends on only $n - m$ and sinc function is an even function. In addition, $\mathbf{R}$ generally has a rank less than N. Therefore, this model can not only be used to simulate radar disturbance samples but also makes ground truth covariance available.

Data from the L-band data set of KASSPER program is the other data set used in our experiments. Note, the KASSPER dataset also makes the true ground truth covariance available and we picked range bins such that their covariance matrices were exactly or approximately Toeplitz. The L-band data set consists of a data cube of 1000 range bins corresponding to the returns from a single coherent processing interval from 11 channels and 32 pulses. Therefore, the dimension of observations (or the spatio temporal product) N is $11 \times 32 = 352$. Other key parameters are detailed in Table 2.1.

We compare the following six different covariance estimation techniques: A host of competing techniques like FML, Eigen-canceller, and shrinkage estimators have been compared with the RCML method in [70]. The results of [70] demonstrate that RCML ouperforms these techniques under all conditions of training data support and hence they are not reproduced here.

- **Sample Covariance Matrix:** The sample covariance matrix is given by $\mathbf{S} = \frac{1}{K}\mathbf{Z}\mathbf{Z}^H$. It is well known that the sample covariance is the unconstrained maximum likelihood estimator under Gaussian disturbance statistics. We refer to the use of this technique as SMI.

- **Iterated Toeplitz Approximation Method:** The iterated Toeplitz approximation method (ITAM) [72] alternatively estimates a rank deficient matrix using the eigenvalue decomposition and then make the resulting matrix Toeplitz by substituting diagonal entries with the average

Table 2.1: KASSPER Dataset-1 parameters

Parameter	Value
Carrier Frequency	1240 MHz
Bandwidth (BW)	10 MHz
Number of Antenna Elements	11
Number of Pulses	32
Pulse Repetition Frequency	1984 Hz
1000 Range Bins	35 km to 50 km
91 Azimuth Angles	87°, 89°, … 267°
128 Doppler Frequencies	-992 Hz, -976.38 Hz, …, 992 Hz
Clutter Power	40 dB
Number of Targets	226 (200 detectable targets)
Range of Target Dop. Freq.	-99.2 Hz to 372 Hz

value of themselves for each diagonal of the estimated matrix. After that, the same process is repeated until the estimated Toeplitz matrix has a desired rank. The estimated covariance satisfies both a desired rank and Toeplitz property and it is closer to the true covariance matrix in the sense of Frobenius norm than the sample covariance matrix.

- **Asymptotic Maximum Likelihood:** The asymptotic maximum likelihood (AML) [19] exploits Toeplitz property of the structured covariance matrix. The authors derived a closed-form formula for Toeplitz covariance matrix estimation and it facilitates computationally efficient implementation. However, they assumed a large number of training samples and their closed-form solution is asymptotically valid. That is, in the low/realistic training regime, estimation performance invariably suffers.

- **Rank Constrained ML estimator:** The RCML estimator [70] has been recently proposed and exploits the clutter rank information of the structured covariance matrix but not Toeplitz property. It is also the first step of the closed form solution of our proposed method.

- **Sequential Quadratic Programming:** Al-Homidan proposed a sequential quadratic programming (SQP) algorithm to find the nearest symmetric positive semi-definite Toeplitz matrix to given a matrix [59]. There are many other Toeplitz approximation algorithms in applied mathematics [60, 61, 62]. We choose the SQP algorithm largely because it guarantees a global minima in approximation error and the l_1 SQP method is considerably faster [59] than alternatives. In practice, the estimator is developed by making a Toeplitz approximation to the RCML estimator. This makes the technique analogous to our proposal of decoupling the rank and Toeplitz constraints in Section 2.3. However, using applied math approximations in a 'black-box' manner has two major drawbacks: 1.) the approximation may not necessarily preserve rank and radar STAP specific structure (e.g. eigenvector matrix is perturbed as well), and 2.) the techniques are numerically involved particularly with an increase in data dimension.

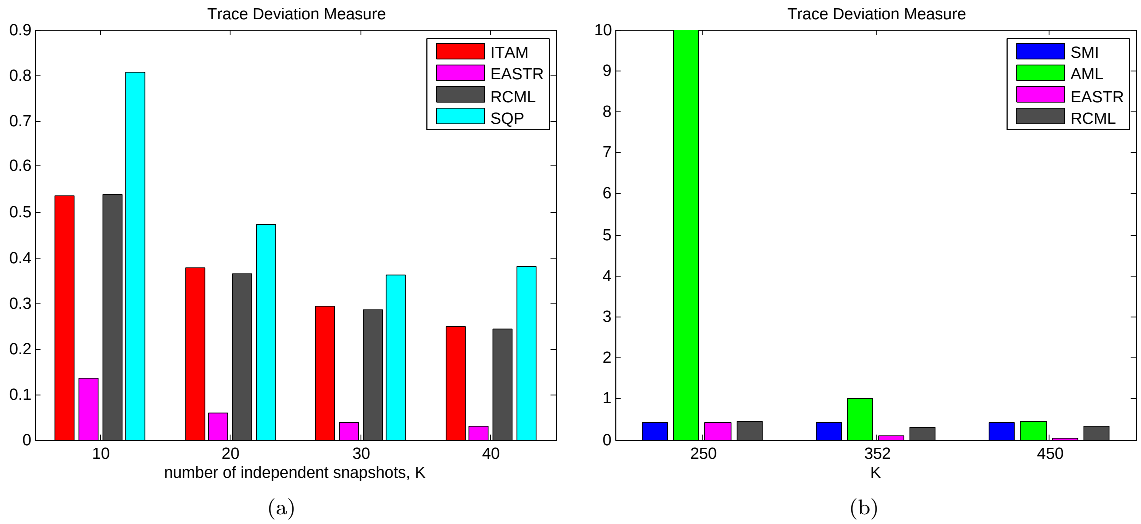


Figure 2.1: Trace deviation measure vs. the number of independent snapshots for (a) simulation model and (b) KASSPER data set.

- **EASTR:** The proposed Efficient Approximation of Structured covariance under joint Toeplitz and Rank (EASTR) constraints. It incorporates Toeplitz structure, the rank of the clutter component as well as the STAP structural constraint.

In the results to follow, the ITAM, RCML, SQP and EASTR exploit rank information. The clutter rank for the simulation model covariance is of course known and for the KASSPER data set was inferred via the Brennan rule.

2.4.2 Whiteness Test

Before using popular radar STAP measures, we apply a 'whiteness test'. The trace deviation measure [56] is one way of evaluating covariance matrix estimators since it captures the extent to which the estimated covariance matrix whitens the true covariance matrix. It is given by

$$TRD(\hat{\mathbf{R}}) = |tr\{\mathbf{R}^{-1}\hat{\mathbf{R}}\}/N - 1| \tag{2.26}$$

Intuitively, we can see that its lower bound is zero when $\hat{\mathbf{R}} = \mathbf{R}$ and smaller value of TRD means better performance.

Fig. 2.1 shows bargraphs of the performance of compared methods for simulation model and KASSPER data set respectively. Fig. 2.1a shows bargraphs of the performance in terms of TRD measure versus the number of training samples. Because the SMI and the AML show very high TRD values, we do not plot them in Fig. 2.1a. Fig. 2.1b similarly shows the result of TRD measure across three training regimes for the KASSPER data set. The TRD measure results in Figs. 2.1a and 2.1b

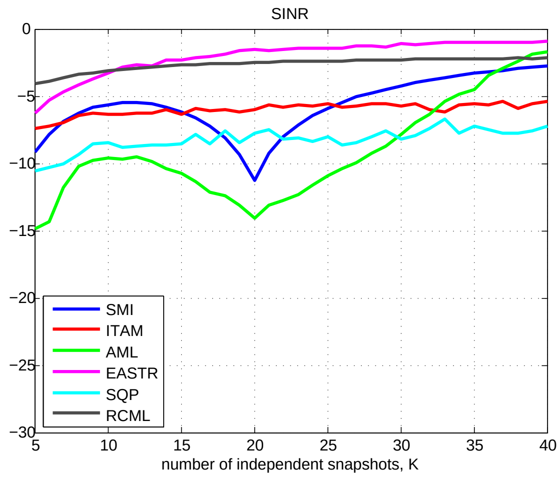


Figure 2.2: Normalized SINR versus the number of independent snapshots for simulation model. $N = 20$

reveal hence that EASTR is infact "structurally" the closest to the true covariance matrix.

2.4.3 Normalized SINR

The normalized SINR measure [44] is commonly used in the radar literature and is given by

$$\eta_i = \frac{|\mathbf{s}^H \hat{\mathbf{R}}_i^{-1} \mathbf{s}|^2}{|\mathbf{s}^H \hat{\mathbf{R}}_i^{-1} \mathbf{R} \hat{\mathbf{R}}_i^{-1} \mathbf{s}||\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}|} \tag{2.27}$$

where $\mathbf{s}$ is the spatio-temporal steering vector, $\hat{\mathbf{R}}_i$ is the data-dependent estimate of $\mathbf{R}$ at the i-th trial, and $\mathbf{R}$ is the true covariance matrix. It is easily seen that $0 < \eta < 1$ and $\eta = 1$ if and only if $\hat{\mathbf{R}} = \mathbf{R}$. The SINR is plotted in dB, that is, $\text{SINR}_i(\text{dB}) = 10 \log_{10} \eta_i$. Therefore, $\text{SINR}_i(\text{dB}) \leq 0$.

We plot the normalized average SINR versus the number of training samples K in Fig. 2.2. In this case, we consider the presense of wideband jamming $J = 3$. In particular, the fractional bandwidth $\beta_i = [0.2, 0, 0.3]$, the powers and phases of jammers are 10 dB, 20 dB, 30 dB and 20 deg, 40 deg, and 60 deg, respectively. When $K < N$ the sample covariance is singular, therefore we used its pseudo-inverse instead of inverse itself.[1] Interpreting the results in Fig. 2.2, it is useful to start with AML which does particularly well when training is generous $K >> N$. However, because AML is asymptotically based - its performance is poor when $K < N$ or $K \approx N$ (K in the vicinity of N is often considered realistic training). Even, the SMI and SQP estimators are better than AML when training is low/realistic. ITAM is effective in very low training as expected because its exploits both rank and Toeplitz constraints (though in a largely heuristic way) - ITAM does not exhibit scalable improvements as training support is increased. EASTR performs the best overall, even better than RCML (which was recently demonstrated to be the most competitive radar STAP estimator [70]) by

[1]Note also that SMI and AML have a dip when $K = 20$ due to numerical instabilities in the $K = N$ training regime. In contrast, ITAM, RCML, EASTR, and SQP guarantee nonsingularity in all training regimes.

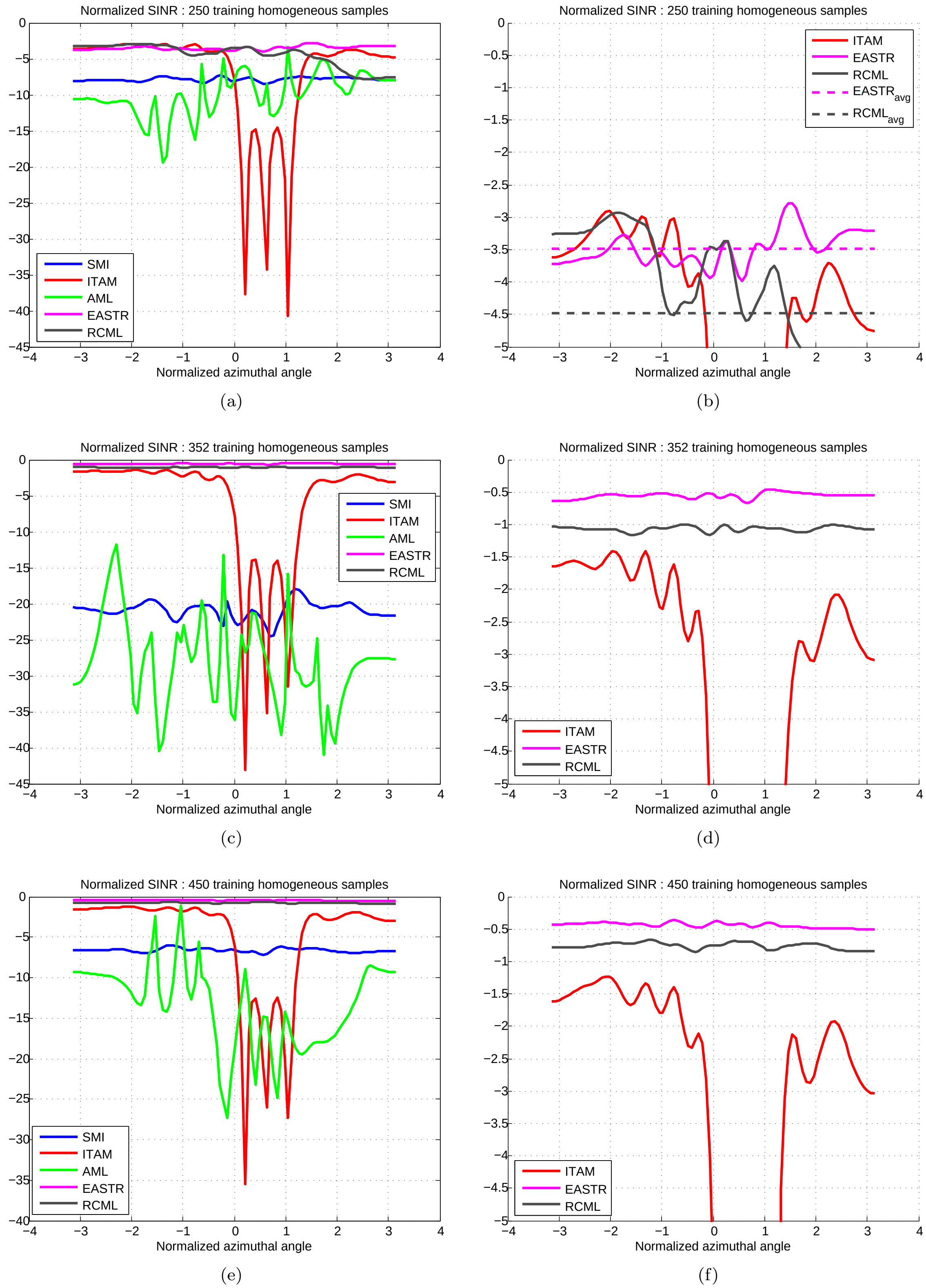
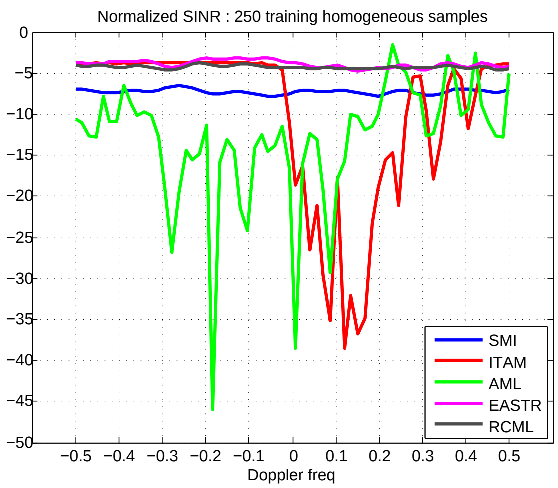
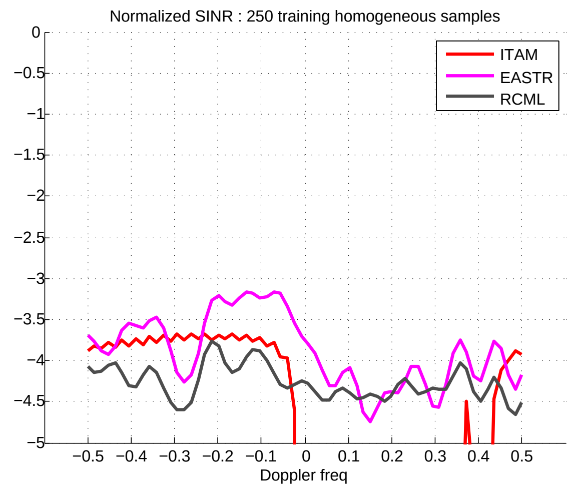


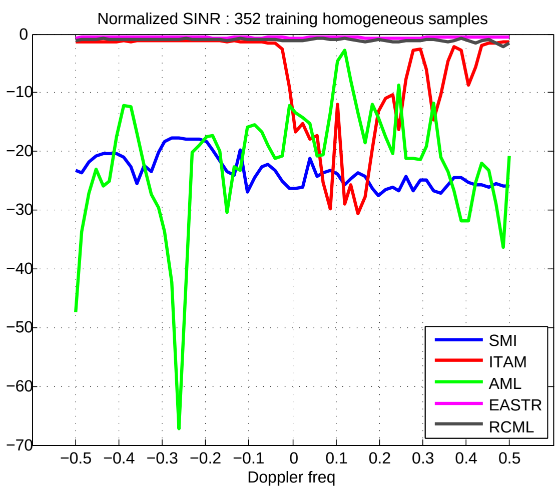
Figure 2.3: Normalized SINR versus azimuthal angle for KASSPER data set. (a) and (b) for $K(=250) < N$, (c) and (d) for $K = N = 352$, and (e) and (f) for $K(=450) > N$



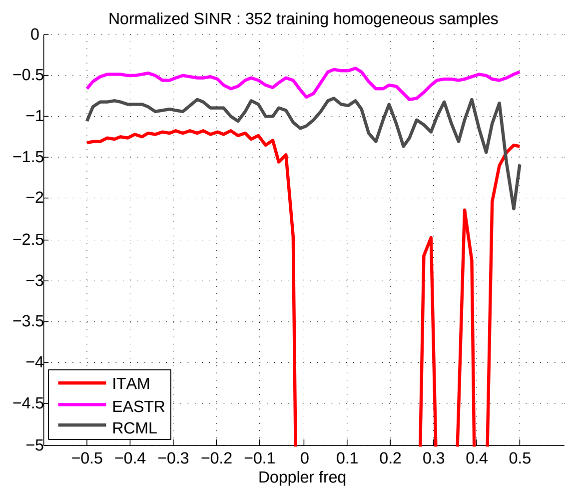
(a)



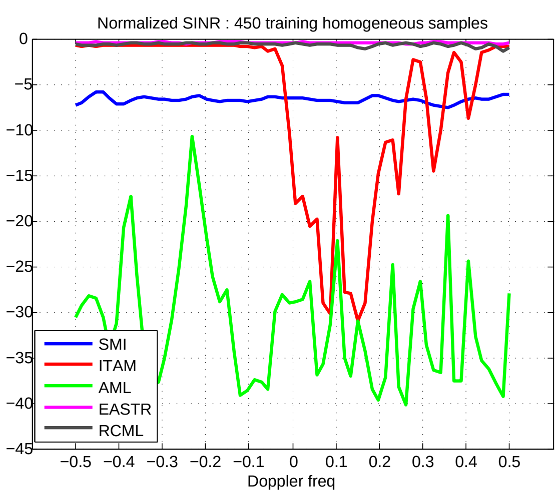
(b)



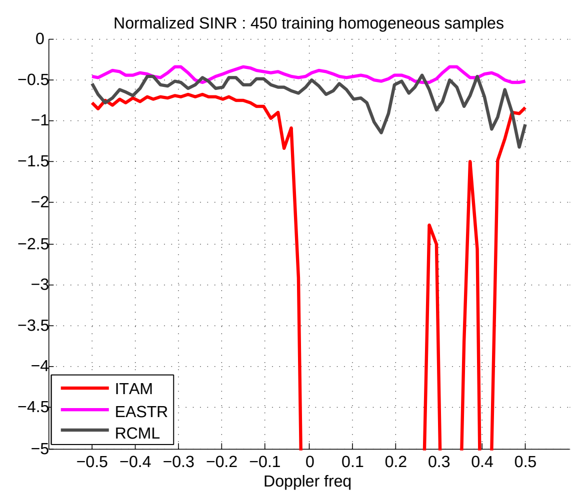
(c)



(d)

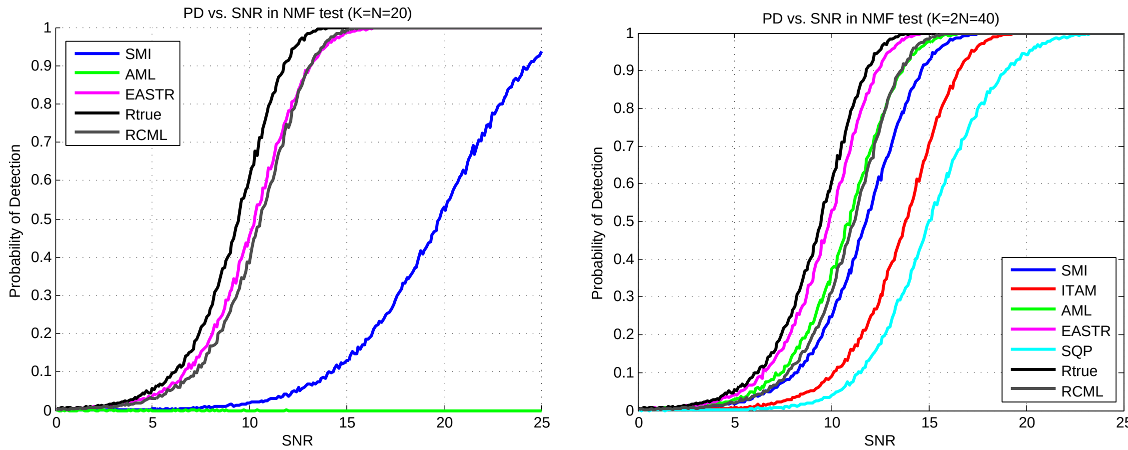


(e)



(f)

Figure 2.4: Normalized SINR versus Doppler frequency for KASSPER data set. (a) and (b) for $K(=250) < N$, (c) and (d) for $K = N = 352$, and (e) and (f) for $K(=450) > N$



(a) (b)

Figure 2.5: Probability of detection vs. SNR for simulation model via normalized matched filter (NMF) test. (a) $K = N = 20$ and (b) $K = 2N = 40$.

virtue of additionally capturing the Toepltiz structure on covariance.

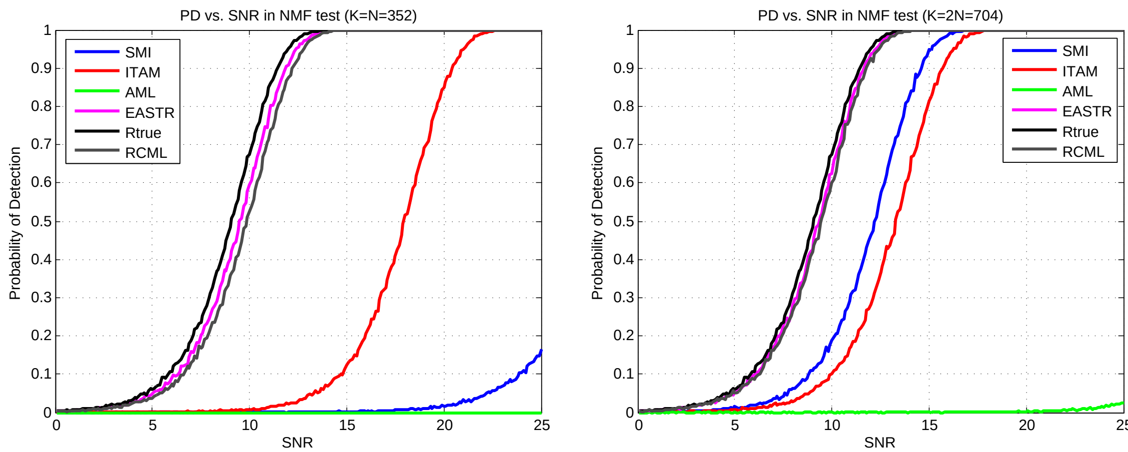
Fig. 2.3 and Fig. 2.4 similarly plots the normalized SINR result for KASSPER data set as a function of the azimuthal angle and Doppler frequency for three different training regimes. Specifically, the first rows of Fig. 2.3 and Fig. 2.4 are corresponding to $K = 250 (< N)$, the second rows are corresponding to $K = N = 352$, and finally the third rows are corresponding to $K > N = 450$ training samples. Plots in the right column show zommed in versions of ITAM, EASTR, and RCML. The sample covariance technique and the AML suffer tremendously when $K \leq N$. For low training, ITAM shows comparable performance to the EASTR and the RCML estimators in some ranges of the azimuthal angle but is worse in some other ranges. On an average (over azimuthal angle and Doppler frequency), EASTR is easily the best in Fig. 2.3 and Fig. 2.4, even providing appreciably gains over the second best RCML estimator. Further, EASTR is stable and effective across all training regimes $K < N, K \approx N$ and $K > N$.

2.4.4 Probability of Detection vs. SNR

In order to compute probability of detection, P_d, we apply the normalized matched filter (NMF) [75] as the test statistic

$$\frac{|\mathbf{s}^H \mathbf{R}^{-1} \mathbf{x}|^2}{[\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}][\mathbf{x}^H \mathbf{R}^{-1} \mathbf{x}]} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{\text{NMF}} \tag{2.28}$$

where $\mathbf{x}$ and K are the observation vector and the number of training samples, respectively. The detection probability P_d is defined as the probability that the value of test statistic is grater than a threshold conditioned on the hypothesis that the received data includes target information. Therefore, it depends on signal to noise ratio (SNR, by virtue of $\mathbf{s}$,) and the estimated covariance matrix. Since P_d



(a) (b)

Figure 2.6: Probability of detection vs. SNR for KASSPER data set via normalized matched filter (NMF) test. (a) $K = N = 352$ and $K = 2N = 704$.

does not typically admit a closed form, we first generate a number of samples from the true covariance to determine λ corresponding to the fixed false alarm rate and then employ Monte Carlo simulations to evaluate P_d corresponding to each estimator. We set a constant false alarm rate to 10^{-4}.

Fig. 2.5 shows the detection probability P_d for simulation model plotted as a function of SNR for different estimators. We use $K = N = 20$ and $K = 2N = 40$ training samples to estimate the covariance matrix in Fig. 2.5a and Fig. 2.5b, respectively. It is well-known that $K = 2N$ training samples are needed to keep the performance within 3 dB. Indeed, we see that the sample covariance matrix has about 3 dB loss vs. the true covariance matrix in Fig. 2.5b. The proposed EASTR is the closest to the P_d achieved by using the true covariance matrix (upper bound) for both cases. In Fig. 2.5a, we do not plot for ITAM and SQP because they do not guarantee positive semi-definiteness of final estimate in the case of $K = N = 20$, so we cannot calculate the detection probabilities for them.

Fig. 2.6 also shows the probability of detection versus SNR plots. We use the same training regimes as used in Section 2.4.3. Fig. 2.6a and Fig. 2.6b plot results for $K = 352$ and $K = 2N = 704$, respectively. We can see similar trends in Fig. 2.6 hence for KASSPER data to the ones for the simulation model in Fig. 2.5. EASTR exhibits the best performance in both plots.

2.5 Conclusions

Our work focuses on jointly exploiting a Toeplitz structure as well as a rank constraint on the clutter covariance for radar STAP. The problem is inherently hard because it is well known that there is no closed form solution for ML estimation under Toeplitz constraint for all training regimes. While past work has provided iterative often expensive solutions, we develop a new estimator that is based on a cascade of two closed forms. The first closed form is the recently proposed RCML estimator. Our core contribution, the second step of Toeplitz approximation performs constrained optimization of eigenvalues to either exactly or approximately satisfy the Toeplitz constraint without compromising the rank. Crucially, this optimization also has a closed form making the overall estimator very friendly from a computational standpoint. Via performance analysis evaluating probability of detection, normalized SINR, and trace deviation measure, our estimator is shown to outperform traditional efforts in Toeplitz and low rank covariance estimation including those based on expensive numerical solutions. Recently, the optimality of the fast maximum likelihood [24] covariance estimator has been proven with respect to cost functions involving the Frobenius or the spectral norm [76]. EASTR can also be investigated for similar notions of optimality. In addition, more analysis of our estimator such as asymptotic convergence can be performed. Finally, practical evaluation may be performed on other radar data sets involving departures from idealized scenarios.

Chapter 3

Robust Covariance Estimation under Imperfect Constraints using Expected Likelihood Approach

3.1 Introduction

Radar systems using multiple antenna elements and processing multiple pulses are widely used in modern radar signal processing since it helps overcome the directivity and resolution limits of a single sensor. Joint adaptive processing in the spatial and temporal domains for the radar systems, called space time adaptive processing (STAP) [47, 48, 44], enables to suppress interfering signals as well as to preserve gain on the desired signal. Interference statistics, in particular the covariance matrix of the disturbance, which must be estimated from secondary training samples in practice plays a critical role on success of STAP. To obtain a reliable estimate of the disturbance covariance matrix, a large number of homogeneous training samples are necessary. This gives rise to a compelling challenge for radar STAP because such generous homogeneous (target free) training is generally not available in practice [7].

Much recent research for radar STAP has been developed to overcome this practical limitation of generous homogeneous training. Specifically, the knowledge-based processing which uses *a priori* information about the interference environment is widely referred in the literature [29, 40] and has merit in the regime of limited training data. These techniques include intelligent training selection [29] and the spatio-temporal degrees of freedom reduction [40, 12, 6]. In addition, covariance matrix estimation techniques the enforce and exploit a particular structure have been pursed as one approach of these techniques. Examples of structure include persymmetry [17], Toeplitz structure [19, 18, 20], circulant structure [21], and eigenstructure [24, 70, 25]. In particular, the fast maximum likelihood

(FML) method [24] which enforces a special eigenstructure that the disturbance covariance matrix represents a scaled identity matrix plus a rank deficient and positive semidefinite clutter component also falls in this category and is shown to be the most competitive technique experimentally.

Recently, the works by Kang *et al.* [70] and Aubry *et al.* [25] have also improved upon the FML by exploiting practical constraints inspired by physical radar environment, specifically the eigenstructure of the disturbance covariance matrix for radar STAP. They employed a rank of the clutter subspace and a condition number of the interference covariance matrix respectively as a constraint as well as the structural constraint used in the FML into the optimization problem. For both methods, though the initial optimization problems are non-convex, the estimation problems are reduced to a convex optimization problems and admit closed-form solutions. Their methods have also been shown to enable higher normalized SINR over the state-of-the art alternatives for the simulation model and the knowledge-aided sensor signal processing and expert reasoning (KASSPER) data set.

In [70], the authors assume the rank of the clutter is given by Brennan rule [1] under ideal conditions of no coupling. However, in practice (under non-ideal conditions) the clutter rank departs from the Brennan rule prediction due to antenna errors and internal clutter motion. In this case, the rank is not known precisely and needs to be determined before using with the RCML estimator. Determination of the number of signals in a measurement record is a classical eigenvalue problem, which has received considerable attention in the past 60 years. It is important to note that the problem does not have a simple and unique solution. Consequently, a number of techniques have been developed to address this problem [77, 78, 79, 80, 81]. In addition, the noise level and the condition number should be estimated as well if they are unknown or non precisely known in practice.

Expected likelihood (EL) approach [82] has been proposed to determine a regularization parameter based on the statistical invariance property of the likelihood ratio (LR) values. More specifically, the probability distribution function (pdf) of LR values for the true covariance matrix depends on only the number of training samples (K) and the dimension of the true covariance matrix (N), not the true covariance itself under a Gaussian assumption on the observations. This statistical independence of LR values on the true covariance itself enables pre-calculation of LR values even though the true covariance is unknown. Finally, the regularization parameters are selected so that the LR value of the estimate agrees as closely as possible with the *median* LR value determined via its pre-characterized pdf.

Contributions: In view of the aforementioned observations, we develop covariance estimation methods which automatically and adaptively determines the values of practical constraints via an expected likelihood approach for practical radar STAP. Our main contributions are:

- **A method of choice of constraints using the EL approach:** We propose a method of a choice of practical constraints employed in the optimization problems for covariance estimation in radar STAP using the expected likelihood approach. The proposed method guides the selection of the constraints via the expected likelihood criteria in the case that the knowledge of the constraints is imperfectly known in practice. We consider three different cases of the constraints in this chapter: 1) only the clutter rank constraint, 2) both the clutter rank and the noise power constraints, and 3) the condition number constraint.

- **Analytical results with formal proofs for three different cases of imperfect constraints:** For each case mentioned above, we develop significant analytical results. We first formally prove that the rank selection problem based on the expected likelihood approach has a unique solution. This guarantees there is only one rank which is the best (global optimal) rank in the sense of the EL approach. Second, we derive a closed form solution of the optimal noise power in the sense of the EL approach for a given rank. This means we do not need iterative or numerical method to find the optimal noise power and enables fast implementation. Finally, we also prove there exists the unique condition number for the condition number selection method via the EL approach.

- **Experimental Results through simulated model and the KASSPER data set:** Experimental investigation on a simulation model and on the KASSPER data set shows that the proposed methods for three different cases outperform alternatives such as the FML, leading rank selection methods in radar literature and statistics, and the ML estimation of the condition number constraint in the sense of normalized SINR.

The rest of the chapter is organized as follows. We provide the proposed methods of the constraint selection problems via the EL approach in Section 3.2. Experimental validation of our method is provided in Section 3.3 wherein we report the performance of the proposed method and compare it against existing methods in terms of normalized SINR on both the simulation model and the KASSPER data set.

3.2 Constraints selection method via Expected Likelihood Approach

3.2.1 Imperfect rank constraint

In Chapter 1, we discuss that the RCML estimator is not only powerful in practice but also computationally cheap and the EL approach is shown to be useful to select parameters so that the estimate is consistent with the true covariance matrix in the sense of the LR value. From Eq. (2.5), we see the RCML solution is a function of the rank r and d_i's which are given in the problem. We propose to use the EL approach to refine and find the *optimal* rank when the rank determined by underlying physics is not necessarily accurate.

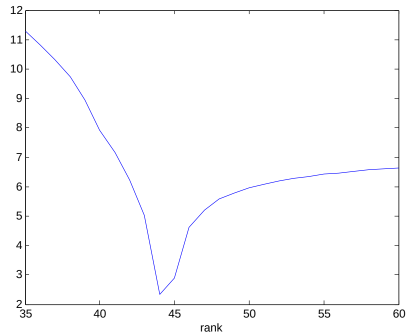


Figure 3.1: $\left(\log\left(\text{LR}\left(\mathbf{R}_{\text{RCML}}(r), \mathbf{Z}\right) / \text{LR}_0\right)\right)^2$ versus r for KASSPER dataset ($K = 2N = 704$)

Now we set up the optimization criterion to find the rank via the EL approach. We find a rank which makes its corresponding LR value closer to LR_0 than any other ranks. That is,

$$\hat{\mathbf{R}}_{\text{RCML}_{\text{EL}}} = \sigma^2 \mathbf{V} \mathbf{\Lambda}^{\star -1}(\hat{r}) \mathbf{V}^H \tag{3.1}$$

where

$$\hat{r} \equiv \arg\min_{r \in \mathbb{Z}} \left| \text{LR}\left(\mathbf{R}_{\text{RCML}}(r), \mathbf{Z}\right) - \text{LR}_0 \right|^2 \tag{3.2}$$

and $\text{LR}\left(\mathbf{R}_{\text{RCML}}(r), \mathbf{Z}\right)$ is given by Eq. (3.3).

Now we investigate the optimization problem (3.2) for the rank selection. Since the eigenvectors of $\mathbf{R}_{\text{RCML}}$ are identical to those of the sample covariance matrix $\mathbf{S}$, the LR value of $\mathbf{R}_{\text{RCML}}$ in Eq. (3.2) can be reduced to the function of the eigenvalues of $\mathbf{R}_{\text{RCML}}$ and $\mathbf{S}$. Let the eigenvalues of $\mathbf{R}_{\text{RCML}}$ and $\mathbf{S}$ be λ_i and d_i (arranged in descending order). Then the LR value of $\mathbf{R}_{\text{RCML}}$ can be simplified to a function of ratio of d_i to λ_i, $\dfrac{d_i}{\lambda_i}$. That is,

$$\begin{aligned} \text{LR}\left(\mathbf{R}_{\text{RCML}}(r), \mathbf{Z}\right) &= \frac{|\hat{\mathbf{R}}_{\text{RCML}}^{-1}(r)\mathbf{S}| \exp N}{\exp\left(\text{tr}\left[\hat{\mathbf{R}}_{\text{RCML}}^{-1}(r)\mathbf{S}\right]\right)} &(3.3) \\ &= \frac{\prod_{i=1}^{N} \frac{d_i}{\lambda_i} \cdot \exp N}{\exp\left[\sum_{i=1}^{N} \frac{d_i}{\lambda_i}\right]} &(3.4) \end{aligned}$$

Lemma 1. *The LR value of the RCML estimator,* $\text{LR}\left(\mathbf{R}_{RCML}(r), \mathbf{Z}\right)$*, is a monotonically increasing function with respect to the rank* r *and there is only one unique* $\hat{r}$ *in the optimization problem* (3.2)*.*

Proof. First, let r be the largest i such that $d_{i+1} \geq \sigma^2$. Then, from the closed form solution of the RCML estimator, the eigenvalues of the RCML estimator with rank i and $i+1$ for given $i < r$ will be

- $\hat{\mathbf{R}}_{\text{RCML}}(i) : d_1, \ d_2, \ \ldots, \ d_i, \ \sigma^2, \ \ldots, \ \sigma^2$

- $\hat{\mathbf{R}}_{\text{RCML}}(i+1):\ d_1,\ \ d_2,\ \ \ldots,\ \ d_i,\ \ d_{i+1},\ \ \sigma^2,\ \ \ldots,\ \ \sigma^2$

since $d_{i+1} \geq \sigma^2$. Then $\dfrac{d_i}{\lambda_i}$ should be

- $\hat{\mathbf{R}}_{\text{RCML}}(i):\ 1,\ \ 1,\ \ \ldots,\ \ 1_i,\ \ \dfrac{d_{i+1}}{\sigma^2},\ \ \ldots,\ \ \dfrac{d_N}{\sigma^2}$
- $\hat{\mathbf{R}}_{\text{RCML}}(i+1):\ 1,\ \ 1,\ \ \ldots,\ \ 1_i,\ \ 1_{i+1},\ \ \dfrac{d_{i+2}}{\sigma^2},\ \ \ldots,\ \ \dfrac{d_N}{\sigma^2}$

From Eq. (3.4), the LR values of the RCML estimators with the ranks i and $i+1$ are

$$\mathrm{LR}(i) = \frac{\dfrac{\exp N}{\sigma^{2(N-i)}} \displaystyle\prod_{k=i+1}^{N} d_k}{\exp(i + \dfrac{1}{\sigma^2} \displaystyle\sum_{k=i+1}^{N} d_k)} \tag{3.5}$$

$$\mathrm{LR}(i+1) = \frac{\dfrac{\exp N}{\sigma^{2(N-i-1)}} \displaystyle\prod_{k=i+2}^{N} d_k}{\exp(i + 1 + \dfrac{1}{\sigma^2} \displaystyle\sum_{k=i+2}^{N} d_k)} \tag{3.6}$$

From Eq. (3.5) and Eq. (3.6), we obtain

$$\begin{aligned}
\mathrm{LR}(i+1) &= \frac{\dfrac{\exp N}{\sigma^{2(N-i-1)}} \displaystyle\prod_{k=i+2}^{N} d_k}{\exp(i + 1 + \dfrac{1}{\sigma^2} \displaystyle\sum_{k=i+2}^{N} d_k)} && (3.7)\\
&= \frac{\dfrac{\exp N}{\sigma^{2(N-i)}} \displaystyle\prod_{k=i+1}^{N} d_k \cdot \dfrac{\sigma^2}{d_{i+1}}}{\exp(i + \dfrac{1}{\sigma^2} \displaystyle\sum_{k=i+1}^{N} d_k) \exp(1 - \dfrac{d_{i+1}}{\sigma^2})} && (3.8)\\
&= \mathrm{LR}(i) \cdot \frac{\sigma^2}{d_{i+1}} \cdot \exp(\frac{d_{i+1}}{\sigma^2} - 1) && (3.9)
\end{aligned}$$

Eq. (3.9) tells us $\mathrm{LR}(i+1)$ can be calculated by multiplying $\mathrm{LR}(i)$ by the coefficient $\dfrac{\sigma^2}{d_{i+1}} \cdot \exp(\dfrac{d_{i+1}}{\sigma^2} - 1)$. Fig. 3.2 shows that

$$\frac{\sigma^2}{d_{i+1}} \cdot \exp(\frac{d_{i+1}}{\sigma^2} - 1) \geq 1 \tag{3.10}$$

for all values of $\dfrac{\sigma^2}{d_{i+1}}$. Therefore, it is obvious that

$$\mathrm{LR}(i+1) \geq \mathrm{LR}(i), \tag{3.11}$$

which means the LR value monotonically increases with respect to i.

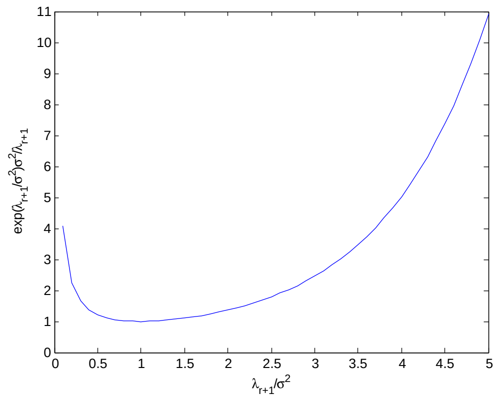


Figure 3.2: The value of the coefficient $\dfrac{\sigma^2}{\lambda_{r+1}} \cdot \exp(\dfrac{\lambda_{r+1}}{\sigma^2} - 1)$

Now, let's consider the other case, $i \geq r$. In this case, since $d_{i+1} < \sigma^2$, it is easily shown that

$$\mathbf{R}_{\text{RCML}}(i) = \mathbf{R}_{\text{RCML}}(i+1) \tag{3.12}$$

Therefore,

$$\text{LR}(i+1) = \text{LR}(i) \tag{3.13}$$

This proves that $\text{LR}(i)$ monotonically increases for all $1 \leq i \leq N$.

□

Lemma 1 gives us a significant analytical result that is the EL approach leads to a unique value of the rank, i.e., when searching over the various values of the rank it is impossible to come up with multiple choices. That also means that it is guaranteed that we can always find the global optimum of r not local optima (minima) for the optimization problem (3.2) regardless of an initial value of r. We plot the values of $\left(\log \left(\text{LR} \left(\mathbf{R}_{\text{RCML}}(r), \mathbf{Z} \right) / \, \text{LR}_0 \right) \right)^2$ versus the rank r for one realization for the KASSPER dataset ($K = 2N = 704$) in Fig. 3.1. Since the LR values are too small in this case, we use a log scale and the ratio between two instead of the distance to see the variation clearly. Note that monotonic increase of the value of $\text{LR} \left(\mathbf{R}_{\text{RCML}}(r), \mathbf{Z} \right)$ w.r.t r guarantees a unique optimal rank even if the optimization function as defined in (3.2) is not necessarily convex in r.

The algorithm to find the optimal rank is simple and not computationally expensive due to the analytical results above. For a given initial rank such as Brennan rule for the KASSPER data set and the number of jammers for a simulation model, we first determine a direction of searching and then find the optimal rank. The procedure of finding the optimal rank is shown in Algorithm 1 in detail.

Algorithm 1 The proposed algorithm to select the rank via EL

Initialize the rank r by physical environment such as Brennan rule.

Evaluate $\mathrm{LR}(r-1)$, $\mathrm{LR}(r)$, $\mathrm{LR}(r+1))$, the LR values of RCML estimators for the ranks $r-1$, r, $r+1$, respectively.

- if $|\,\mathrm{LR}(r+1) - \mathrm{LR}_0\,| < |\,\mathrm{LR}(r) - \mathrm{LR}_0\,|$
  $\rightarrow$ increase r by 1 until $|\,\mathrm{LR}(r) - \mathrm{LR}_0\,|$ is minimized to find $\hat{r}$.
- elseif $|\,\mathrm{LR}(r-1) - \mathrm{LR}_0\,| < |\,\mathrm{LR}(r) - \mathrm{LR}_0\,|$
  $\rightarrow$ decrease r by 1 until $|\,\mathrm{LR}(r) - \mathrm{LR}_0\,|$ is minimized to find $\hat{r}$.
- else $\hat{r} = r$, the initial rank.

3.2.2 Imperfect rank and noise power constraints

In this section, we investigate the second case that both the rank r and the noise power σ^2 are not perfectly known. We propose the estimation of both the rank and the noise level based on the EL approach. The estimator with both the rank and the noise power obtained by the EL approach is given by

$$\hat{\mathbf{R}}_{\mathrm{RCML_{EL}}} = \hat{\sigma}^2 \mathbf{V} \mathbf{\Lambda}^{\star -1}(\hat{r}) \mathbf{V}^H \tag{3.14}$$

where

$$(\hat{r}, \hat{\sigma}^2) \equiv \arg \min_{r \in \mathbb{Z}, \sigma^2 > 0} \left| \mathrm{LR}\left(\mathbf{R}_{\mathrm{RCML}}(r, \sigma^2), \mathbf{Z}\right) - \mathrm{LR}_0 \right|^2 \tag{3.15}$$

In section 3.2.1, we have shown that the optimal rank via the EL approach is uniquely obtained for a fixed σ^2. Now we analyze the LR values of the RCML estimator for various σ^2 and a fixed rank.

Lemma 2. *For a fixed rank, the LR value of the RCML estimator, which is a function of σ^2, has a maximum value at $\sigma^2 = \sigma^2_{ML}$. It monotonically increases for $\sigma^2 < \sigma^2_{ML}$ and monotonically decreases for $\sigma^2 > \sigma^2_{ML}$.*

Proof. In this section, I investigate the LR values for varying noise level σ^2 and a given rank r. From Eq. (3.5) we obtain the LR value when the rank is r,

$$\mathrm{LR}(\sigma^2) = \frac{\dfrac{\exp N}{\sigma^{2(N-r)}} \displaystyle\prod_{k=r+1}^{N} d_k}{\exp(r + \dfrac{1}{\sigma^2} \displaystyle\sum_{k=r+1}^{N} d_k)} \tag{3.16}$$

For simplicity, let $\sigma^2 = t$ then Eq. (3.16) can be simplified as

$$\mathrm{LR}(t) = \frac{e^{N-r} \displaystyle\prod_{k=r+1}^{N} d_k}{t^{N-r} e^{\frac{\sum_{k=r+1}^{N} d_k}{t}}} \tag{3.17}$$

Now let $\sum_{k=r+1}^{N} d_k = d_s$ and $\prod_{k=r+1}^{N} d_k = d_p$, then

$$\begin{aligned} \mathrm{LR}(t) &= \frac{e^{N-r} d_p}{t^{N-r} e^{\frac{d_s}{t}}} && (3.18) \\ &= d_p e^{N-r} t^{r-N} e^{-\frac{d_s}{t}} && (3.19) \end{aligned}$$

To analyze increasing or decreasing property Eq. (3.19), I calculate its first derivative. Since $d_p e^{N-r}$ is a positive constant, it does not affect increasing or decreasing of the function. Therefore,

$$\begin{aligned} & (t^{r-N} e^{-\frac{d_s}{t}})' \\ &= (r-N) t^{r-N-1} e^{-d_s/t} + t^{r-N} e^{-d_s/t} \frac{d_s}{t^2} && (3.20) \\ &= (r-N) t^{r-N-1} e^{-d_s/t} + t^{r-N-2} e^{-d_s/t} d_s && (3.21) \\ &= t^{r-N-2} \big((r-N)t + d_s \big) e^{-d_s/t} && (3.22) \end{aligned}$$

Since t^{r-N-2} and $e^{-d_s/t}$ are always positive, the first derivative $(t^{r-N} e^{-\frac{d_s}{t}})' = 0$ if and only if

$$t = \frac{d_s}{N-r} = \frac{\sum_{k=r+1}^{N} d_k}{N-r} \qquad (3.23)$$

and it is positive when $t < \frac{\sum_{k=r+1}^{N} d_k}{N-r}$ and negative otherwise. This means that $\mathrm{LR}(\sigma^2)$ increases for $\sigma^2 < \frac{\sum_{k=r+1}^{N} d_k}{N-r}$ and decreases for $\sigma^2 > \frac{\sum_{k=r+1}^{N} d_k}{N-r}$. The LR value is maximized when $\sigma^2 = \frac{\sum_{k=r+1}^{N} d_k}{N-r}$. Note that $\frac{\sum_{k=r+1}^{N} d_k}{N-r}$ is the average value of $N-r$ smallest eigenvalues of the sample covariance matrix and in fact a maximum likelihood solution of σ^2 as shown in the RCML estimator [70]. □

Fig. 3.3 shows an example of the LR values as a function of the noise level σ^2. As shown in Lemma 2, we see that the LR value is maximized for the ML solution of σ^2 and monotonically increases and decreases for each direction. It is obvious that we have three cases of the solution of the optimal noise power from Lemma 2: 1) no solution, 2) only one solution, and 3) two optimal solution. Now we discuss how to obtain the optimal noise power for a fixed rank.

Lemma 3. *The noise power obtained by the expected likelihood approach, $\hat{\sigma}^2_{EL}$, is given by*

$$\hat{\sigma}^2_{EL} = \exp\left(W_k \left(\frac{b}{a} e^{-\frac{c}{a}} \right) + \frac{c}{a} \right) \qquad (3.24)$$

where $W_k(z)$ is the k-th branch of Lambert W function and

$$\begin{cases} a = r - N \\ b = \sum_{k=r+1}^{N} d_k \\ c = \log \mathrm{LR}_0 - \log\left(\prod_{k=r+1}^{N} d_k\right) + a \end{cases} \tag{3.25}$$

Proof. For a given rank r, the optimal solution of the noise power via the EL approach, $\hat{t}(=\hat{\sigma}^2_{\mathrm{EL}})$, is the solution of $\mathrm{LR}(t) = \mathrm{LR}_0$. From Eq. (3.19), that is, $\hat{t}$ is the solution of the equation given by

$$d_p e^{N-r} t^{r-N} e^{-\frac{d_s}{t}} = \mathrm{LR}_0 \tag{3.26}$$

Taking log on both side leads

$$\log d_p + N - r + (r - N)\log t - \frac{d_s}{t} = \log \mathrm{LR}_0 \tag{3.27}$$

For simplification, we take substitutions of variables,

$$\begin{cases} a = r - N \\ b = \sum_{k=r+1}^{N} d_k \\ c = \log \mathrm{LR}_0 - \log\left(\prod_{k=r+1}^{N} d_k\right) + a \end{cases} \tag{3.28}$$

Then, Eq. (3.27) is simplified to an equation of t,

$$a \log t - \frac{b}{t} = c \tag{3.29}$$

Again, let $u = \log t$. Then, since $t = e^u$, we obtain

$$au - be^{-u} = c \tag{3.30}$$

$$e^{-u} = \frac{a}{b}u - \frac{c}{b} \tag{3.31}$$

Now let $s = u - \frac{c}{a}$. Then, the equation is

$$e^{-s-\frac{c}{a}} = \frac{a}{b}s \tag{3.32}$$

$$se^{s} = \frac{b}{a}e^{-\frac{c}{a}} \tag{3.33}$$

The solution of Eq. (3.33) is known to be obtained using Lambert W function [83]. That is,

$$s = W\left(\frac{b}{a}e^{-\frac{c}{a}}\right) \tag{3.34}$$

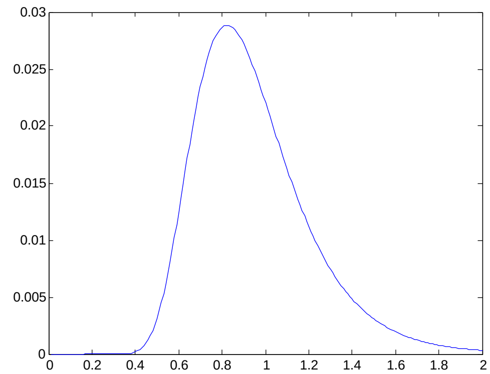


Figure 3.3: The LR value versus σ^2 for the simulation model, $N = 20$, $K = 40$, $r = 5$

where $W(\cdot)$ is a Lambert W function which is defined to be the function satisfying

$$W(z)e^{W(z)} = z \tag{3.35}$$

Finally, we obtain

$$u = W\left(\frac{b}{a}e^{-\frac{c}{a}}\right) + \frac{c}{a} \tag{3.36}$$

and

$$\hat{\sigma}^2_{\mathrm{EL}} = \hat{t} = \exp\left(W\left(\frac{b}{a}e^{-\frac{c}{a}}\right) + \frac{c}{a}\right) \tag{3.37}$$

$\square$

Lemma 3 shows that there is a closed form solution of the optimal noise power for a fixed rank. Therefore we do not need any iterative and numerical algorithms to obtain both the optimal rank and noise power.

Now we propose the method to alternately find the optimal solution of both the rank and the noise power. For a fixed σ^2, we can obtain the optimal rank via Algorithm 1. For a fixed rank, we should consider three cases described above. The first case implies that the LR value corresponding σ^2_{ML} is less than LR_0 and therefore, we increase the rank until the solution of σ^2 exists. In the second case, we can easily determine $\hat{\sigma}^2 = \sigma^2_{\mathrm{ML}}$. For the third case that there are two solutions of σ^2, we have to choose one among two EL solutions and the ML solution. We experimentally observe that the threshold in the test statistics such as the normalized matched filter is typically smaller for the better estimator in the sense of the normalized SINR and the probability of detection from our experiments. Therefore, we choose one of σ^2_{ML}, σ^2_{EL1}, σ^2_{EL2}, which generates the smallest value of the test statistics. The detail procedure of the algorithm is described in Algorithm 2.

Algorithm 2 The proposed algorithm to select the rank and the noise level via EL

1: Initialize the rank r by physical environment such as Brennan rule or the number of jammers.
2: If there is no solution of σ^2 for given r, increase r until the solution of σ^2 exists.
3: Obtain $\sigma^2_{\mathrm{ML}} = \frac{1}{N-r}\sum_{i=r+1}^{N} d_i$.
4: For given σ^2_{ML}, find a new r using Algorithm 1.
5: Repeat Step 3 and Step 4 until the rank r converges.
6: After r is determined, choose $\hat{\sigma}^2$ among σ^2_{ML}, σ^2_{EL1}, σ^2_{EL2}.

3.2.3 Imperfect condition number constraint

Now we discuss the proposed method to determine the condition number constraint through the EL approach in this section. The condition number constrained ML estimator is a function of the condition number $K_{\max}$. Therefore, the final estimate is also a function of $K_{\max}$. Similar to what we have done in previous sections, we find an optimal condition number so that the LR value of the estimated covariance matrix should be same as a statistical median value of the LR value of the true covariance matrix, that is

$$\hat{\mathbf{R}}_{\mathrm{CNCML_{EL}}} = \hat{\sigma}^2 \mathbf{V} \mathbf{\Lambda}^{\star -1}(\hat{K}_{\max})\mathbf{V}^H \tag{3.38}$$

where

$$\hat{K}_{\max} \equiv \arg\min_{\mathrm{K_{max}} \geq 1} \left| \mathrm{LR}\left(\mathbf{R}_{\mathrm{CNCML}}(\mathrm{K_{max}}), \mathbf{Z}\right) - \mathrm{LR}_0 \right|^2 \tag{3.39}$$

Before we discuss the algorithm to find the optimal condition number, we analyze the closed form solution for the condition number constrained ML estimation which is proposed in [25]. We derive a more explicit closed form solution.

Lemma 4. *The more simplified closed form solution of the condition number constrained ML estimator is given by*

1. $d_1 \leq \sigma^2$,

$$\hat{\mathbf{R}}_{CN} = \sigma^2 \mathbf{I} \tag{3.40}$$

2. $\sigma^2 \leq d_1 \leq \sigma^2 K_{\max}$,

$$\hat{\mathbf{R}}_{CN} = \hat{\mathbf{R}}_{FML} \tag{3.41}$$

3. $d_1 > \sigma^2 K_{\max}$ *and* $K_{\max} \geq \frac{\sum_{i=1}^{c} d_i}{c - \sum_{\bar{N}+1}^{N}(d_i - 1)}$,

$$\hat{\mathbf{R}}_{CN} = \mathbf{\Phi}\,\mathrm{diag}(\boldsymbol{\lambda}^*)\mathbf{\Phi}^H \tag{3.42}$$

where

$$\boldsymbol{\lambda}^\star = \left[\sigma^2 K_{\max}, \ldots, \sigma^2 K_{\max}, d_{c+1}, \ldots, d_{\bar{N}}, \sigma^2, \ldots, \sigma^2\right], \tag{3.43}$$

c and $\bar{N}$ are the vector of the eigenvalues of the estimate, the largest indices so that $d_c > \sigma^2 K_{\max}$, and $d_{\bar{N}} \geq \sigma^2$

4. $d_1 > \sigma^2 K_{\max}$ *and* $K_{\max} < \frac{\sum_{i=1}^{c} d_i}{c - \sum_{\bar{N}+1}^{N}(d_i - 1)}$,

$$\boldsymbol{\lambda}^\star = [\frac{\sigma^2}{u}, \ldots, \frac{\sigma^2}{u}, d_{p+1}, \ldots, d_q, \frac{\sigma^2}{uK_{\max}}, \ldots, \frac{\sigma^2}{uK_{\max}}] \tag{3.44}$$

And the condition numbers of the estimates are 1, $\frac{d_1}{\sigma^2}$, $\mathrm{K}_{\max}$, *and* $\mathrm{K}_{\max}$, *respectively.*

Proof. We consider 5 cases provided in [25].

1. $d_1 \leq \sigma^2 \leq \sigma^2\,\mathrm{K}_{\max}$

   Since $u^\star = \frac{1}{\mathrm{K}_{\max}}$,

$$\begin{aligned} \lambda_i^\star &= \min(\min(\mathrm{K}_{\max}\,u^\star, 1), \max(u^\star, \frac{1}{\bar{d}_i})) && (3.45)\\ &= \min(\min(1,1), \max(\frac{1}{\mathrm{K}_{\max}}, \frac{1}{\bar{d}_i})) && (3.46)\\ &= \min(1, \frac{1}{\bar{d}_i}) = 1 && (3.47) \end{aligned}$$

   Therefore,

$$\hat{\mathbf{R}}_{\mathrm{CN}} = \sigma^2\mathbf{I} \tag{3.48}$$

   and the condition number is 1.

2. $\sigma^2 < d_1 \leq \mathrm{K}_{\max}$

   Since $u^\star = \frac{1}{\bar{d}_1}$,

$$\begin{aligned} \lambda_i^\star &= \min(\min(\mathrm{K}_{\max}\,u^\star, 1), \max(u^\star, \frac{1}{\bar{d}_i})) && (3.49)\\ &= \min(\min(\frac{\mathrm{K}_{\max}}{\bar{d}_1}, 1), \max(\frac{1}{\bar{d}_1}, \frac{1}{\bar{d}_i})) && (3.50)\\ &= \min(1, \frac{1}{\bar{d}_i}) && (3.51)\\ &= \begin{cases} \frac{1}{\bar{d}_i} & \bar{d}_i \geq 1 \\ 1 & \bar{d}_i < 1 \end{cases} && (3.52) \end{aligned}$$

   Therefore,

$$\hat{\mathbf{R}}_{\mathrm{CN}} = \hat{\mathbf{R}}_{\mathrm{FML}} \tag{3.53}$$

   and the condition number is $\frac{d_1}{\sigma^2}$.

3. $d_1 > \sigma^2\,\mathrm{K}_{\max}$ and $u^\star = \frac{1}{\bar{d}_1}$

   $\frac{dG(u)}{du}|_{u=\frac{1}{\bar{d}_1}}$ must be zero if $u^\star = \frac{1}{\bar{d}_1}$. The first derivative of $G_i(u)$ is given by

$$G_i'(u) = \begin{cases} -\frac{1}{u} + \mathrm{K}_{\max}\,\bar{d}_i & \text{if} \quad 0 < u \leq \frac{1}{\mathrm{K}_{\max}} \\ 0 & \text{if} \quad \frac{1}{\mathrm{K}_{\max}} \leq u \leq 1 \end{cases} \tag{3.54}$$

for $\bar{d}_i \leq 1$, and

$$G_i'(u) = \begin{cases} -\frac{1}{u} + \mathrm{K_{max}}\,\bar{d}_i & \text{if} \quad 0 < u \leq \frac{1}{\mathrm{K_{max}}\,\bar{d}_i} \\ 0 & \text{if} \quad \frac{1}{\mathrm{K_{max}}\,\bar{d}_i} < u \leq \frac{1}{\bar{d}_i} \\ -\frac{1}{u} + \bar{d}_i & \text{if} \quad \frac{1}{\bar{d}_i} \leq u \leq 1 \end{cases} \tag{3.55}$$

for $\bar{d}_i > 1$. Therefore,

$$\frac{dG(u)}{du}|_{u=\frac{1}{\bar{d}_1}} = \sum_{i=\bar{N}+1}^{N} (\mathrm{K_{max}}\,\bar{d}_i - \bar{d}_1) + \sum_{i=p}^{\bar{N}} (\mathrm{K_{max}}\,\bar{d}_i - \bar{d}_1) \tag{3.56}$$

where p is the greatest index such that $\frac{1}{\bar{d}_1} < \frac{1}{\mathrm{K_{max}}\,\bar{d}_p}$. For $i = \bar{N}, \ldots, N$, since $\bar{d}_i \leq 1$,

$$\mathrm{K_{max}}\,\bar{d}_i - \bar{d}_1 < \mathrm{K_{max}} - \bar{d}_1 < 0 \tag{3.57}$$

and for $i = p, \ldots, \bar{N} - 1$, since $\bar{d}_1 > \mathrm{K_{max}}\,\bar{d}_i$, $\mathrm{K_{max}}\,\bar{d}_i - \bar{d}_1 < 0$. Therefore, in this case, it is obvious that

$$\frac{dG(u)}{du}|_{u=\frac{1}{\bar{d}_1}} < 0 \tag{3.58}$$

4. $d_1 > \sigma^2\,\mathrm{K_{max}}$ and $u^\star = \frac{1}{\mathrm{K_{max}}}$

   Aubry *et al.* [25] showed that $u^\star = \frac{1}{\mathrm{K_{max}}}$ if $\frac{dG(u)}{du}|_{u=\frac{1}{\mathrm{K_{max}}}} \leq 0$. From Eq. (3.54) and Eq. (3.55),

$$\frac{dG(u)}{du}|_{u=\frac{1}{\mathrm{K_{max}}}} = \sum_{i=\bar{N}+1}^{N} \mathrm{K_{max}}(\bar{d}_i - 1) + \sum_{i=1}^{p} (\bar{d}_i - \mathrm{K_{max}}) \tag{3.59}$$

where p is the greatest index such that $\bar{d}_p > \mathrm{K_{max}}$. Therefore,

$$\tfrac{dG(u)}{du}|_{u=\frac{1}{\mathrm{K_{max}}}} \leq 0 \tag{3.60}$$

$$\Leftrightarrow \quad \sum_{i=\bar{N}+1}^{N} \mathrm{K_{max}}(\bar{d}_i - 1) + \sum_{i=1}^{p} (\bar{d}_i - \mathrm{K_{max}}) \leq 0 \tag{3.61}$$

$$\Leftrightarrow \quad \mathrm{K_{max}}(\textstyle\sum_{i=\bar{N}+1}^{N} (\bar{d}_i - 1) - p) + \sum_{i=1}^{p} \bar{d}_i \leq 0 \tag{3.62}$$

$$\Leftrightarrow \quad \mathrm{K_{max}}(\textstyle\sum_{i=\bar{N}+1}^{N} (\bar{d}_i - 1) - p) \leq -\sum_{i=1}^{p} \bar{d}_i \tag{3.63}$$

$$\Leftrightarrow \quad \mathrm{K_{max}} \geq \tfrac{\sum_{i=1}^{p} \bar{d}_i}{p - \sum_{i=\bar{N}+1}^{N} (\bar{d}_i - 1)} \tag{3.64}$$

In this case,

$$
\begin{aligned}
\lambda_i^\star &= \min(\min(\mathrm{K}_{\max} u^\star, 1), \max(u^\star, \frac{1}{\bar{d}_i})) && (3.65)\\
&= \min(\min(1,1), \max(\frac{1}{\mathrm{K}_{\max}}, \frac{1}{\bar{d}_i})) && (3.66)\\
&= \min(1, \max(\frac{1}{\mathrm{K}_{\max}}, \frac{1}{\bar{d}_i})) && (3.67)\\
&= \begin{cases} \min(1, \frac{1}{\mathrm{K}_{\max}}) & \bar{d}_i \geq \mathrm{K}_{\max} \\ \min(1, \frac{1}{\bar{d}_i}) & \bar{d}_i < \mathrm{K}_{\max} \end{cases} && (3.68)\\
&= \begin{cases} \frac{1}{\mathrm{K}_{\max}} & \bar{d}_i \geq \mathrm{K}_{\max} \\ \frac{1}{\bar{d}_i} & \bar{1} \leq \bar{d}_i < \mathrm{K}_{\max} \\ 1 & \bar{d}_i < 1 \end{cases} && (3.69)
\end{aligned}
$$

Finally we obtain

$$\boldsymbol{\lambda}^\star = \left[\sigma^2 K_{\max}, \ldots, \sigma^2 K_{\max}, d_{p+1}, \ldots, d_{\bar{N}}, \sigma^2, \ldots, \sigma^2\right], \tag{3.70}$$

where p and $\bar{N}$ are the largest indices so that $d_p > \sigma^2 K_{\max}$ and $d_{\bar{N}} \geq \sigma^2$, respectively.

5. $d_1 > \sigma^2 K_{\max}$ and $\mathrm{K}_{\max} < \frac{\sum_{i=1}^{p} \bar{d}_i}{p - \sum_{i=\bar{N}+1}^{N} (\bar{d}_i - 1)}$
   In this case, since $\frac{1}{d_1} < u^\star < \frac{1}{\mathrm{K}_{\max}}$,

$$
\begin{aligned}
\lambda_i^\star &= \min(\min(\mathrm{K}_{\max} u^\star, 1), \max(u^\star, \frac{1}{\bar{d}_i})) && (3.71)\\
&= \min(\mathrm{K}_{\max} u^\star, \max(u^\star, \frac{1}{\bar{d}_i})) && (3.72)\\
&= \begin{cases} \min(\mathrm{K}_{\max} u^\star, u^\star) & \bar{d}_i \geq \frac{1}{u^\star} \\ \min(\mathrm{K}_{\max} u^\star, \frac{1}{\bar{d}_i}) & \bar{d}_i < \frac{1}{u^\star} \end{cases} && (3.73)\\
&= \begin{cases} u^\star & \bar{d}_i \geq \frac{1}{u^\star} \\ \frac{1}{\bar{d}_i} & \frac{1}{\mathrm{K}_{\max} u^\star} \leq \bar{d}_i \leq \frac{1}{u^\star} \\ \mathrm{K}_{\max} u^\star & \bar{d}_i < \frac{1}{\mathrm{K}_{\max} u^\star} \end{cases} && (3.74)
\end{aligned}
$$

Therefore, we obtain

$$\boldsymbol{\lambda}^\star = [\frac{\sigma^2}{u^\star}, \ldots, \frac{\sigma^2}{u^\star}, d_{p+1}, \ldots, d_q, \frac{\sigma^2}{u^\star K_{\max}}, \ldots, \frac{\sigma^2}{u^\star K_{\max}}] \tag{3.75}$$

where p and q are the largest indices so that $d_p > \frac{\sigma^2}{u}$ and $d_q > \frac{\sigma^2}{u\,\mathrm{K}_{\max}}$, respectively.

□

From Lemma 4, for the first two cases that is $d_1 \leq \sigma^2 K_{\max}$, the estimator is either a scaled identity matrix or the FML. Therefore, there is no need to find an optimal condition number in these cases since the estimator is not a function of the condition number.

Now we investigate uniqueness of the optimal condition number as we have done in the case of only rank constraint for the last two cases where the optimal eigenvalues are functions of the condition number.

Lemma 5. *The LR value of the condition number ML estimator is a monotonically increasing function with respect to the condition number $K_{\max}$ and there is only one unique $K_{\max_{EL}}$.*

Proof. 1. $d_1 \leq \sigma^2$

$$\hat{\mathbf{R}}_{\mathrm{CN}} = \sigma^2 \mathbf{I} \tag{3.76}$$

In this case, $\hat{\mathbf{R}}_{\mathrm{CN}}$ does not change, so $\mathrm{LR}(\mathrm{K}_{\max})$ is a constant.

2. $\sigma^2 \leq d_1 \leq \sigma^2 K_{\max}$

$$\hat{\mathbf{R}}_{\mathrm{CN}} = \hat{\mathbf{R}}_{\mathrm{FML}} \tag{3.77}$$

In this case, $\hat{\mathbf{R}}_{\mathrm{CN}}$ does not change, so $\mathrm{LR}(\mathrm{K}_{\max})$ is a constant.

3. $d_1 > \sigma^2 K_{\max}$ and $K_{\max} \geq \frac{\sum_{i=1}^{p} d_i}{c - \sum_{\bar{N}+1}^{N}(d_i - 1)}$

$$\hat{\mathbf{R}}_{\mathrm{CN}} = \mathbf{\Phi}\, \mathrm{diag}(\boldsymbol{\lambda}^*) \mathbf{\Phi}^H \tag{3.78}$$

where

$$\boldsymbol{\lambda}^\star = \left[\sigma^2 K_{\max}, \ldots, \sigma^2 K_{\max}, d_{p+1}, \ldots, d_{\bar{N}}, \sigma^2, \ldots, \sigma^2\right], \tag{3.79}$$

p and $\bar{N}$ are the largest indices so that $d_p > \sigma^2 K_{\max}$ and $d_{\bar{N}} \geq \sigma^2$, respectively.

$$\begin{aligned}
&\mathrm{LR}(\mathrm{K}_{\max}) \\
&= \frac{\prod_{i=1}^{N} \frac{d_i}{\lambda_i} e^N}{\exp(\sum_{i=1}^{N} \frac{d_i}{\lambda_i})} \\
&= \frac{\displaystyle\prod_{i=1}^{p} \frac{d_i}{\sigma^2\, \mathrm{K}_{\max}} \cdot \prod_{i=p+1}^{\bar{N}} 1 \cdot \prod_{i=\bar{N}+1}^{N} \frac{d_i}{\sigma^2} \cdot e^N}{\exp(\displaystyle\sum_{i=1}^{p} \frac{d_i}{\sigma^2\, \mathrm{K}_{\max}} + \sum_{i=p+1}^{\bar{N}} 1 + \sum_{i=\bar{N}+1}^{N} \frac{d_i}{\sigma^2})} \\
&= \frac{\prod_{i=1}^{p} \frac{d_i}{\sigma^2\, \mathrm{K}_{\max}} \cdot \prod_{i=\bar{N}+1}^{N} \frac{d_i}{\sigma^2} \cdot e^N}{\exp(\displaystyle\sum_{i=1}^{p} \frac{d_i}{\sigma^2\, \mathrm{K}_{\max}}) \cdot e^{\bar{N}-p} \cdot \exp(\displaystyle\sum_{i=\bar{N}+1}^{N} \frac{d_i}{\sigma^2})}
\end{aligned}$$

(a) within the range where p remains same

$$
\begin{aligned}
&\mathrm{LR}(\mathrm{K}_{\max}) \\
&= \frac{\prod_{i=1}^{p} \frac{d_i}{\sigma^2 \mathrm{K}_{\max}} \cdot \prod_{i=\bar{N}+1}^{N} \frac{d_i}{\sigma^2} \cdot e^N}{\exp(\sum_{i=1}^{p} \frac{d_i}{\sigma^2 \mathrm{K}_{\max}}) \cdot e^{\bar{N}-p} \cdot \exp(\sum_{i=\bar{N}+1}^{N} \frac{d_i}{\sigma^2})} \\
&= c_1 \frac{\prod_{i=1}^{p} \frac{d_i}{\sigma^2 \mathrm{K}_{\max}}}{\exp(\sum_{i=1}^{p} \frac{d_i}{\sigma^2 \mathrm{K}_{\max}})} \\
&= c_1 \frac{\frac{1}{(\sigma^2 \mathrm{K}_{\max})^p} \prod_{i=1}^{p} d_i}{\exp(\frac{1}{\sigma^2 \mathrm{K}_{\max}} \sum_{i=1}^{p} d_i)} \\
&= c_1 \frac{\frac{1}{(\sigma^2 \mathrm{K}_{\max})^p} \prod_{i=1}^{p} d_i}{(\exp(\sum_{i=1}^{p} d_i))^{\frac{1}{\sigma^2 \mathrm{K}_{\max}}}} \\
&= c_2 \frac{(\frac{1}{\mathrm{K}_{\max}})^p}{c_3^{\frac{1}{\mathrm{K}_{\max}}}} \\
&= c_2 \frac{1}{(\mathrm{K}_{\max})^p \cdot c_3^{\frac{1}{\mathrm{K}_{\max}}}}
\end{aligned}
$$

where $c_1 = \frac{\prod_{i=\bar{N}+1}^{N} \frac{d_i}{\sigma^2} \cdot e^N}{\exp(\bar{N}-p) \cdot \exp(\sum_{i=\bar{N}+1}^{N} \frac{d_i}{\sigma^2})}$, $c_2 = c_1 \frac{\prod_{i=1}^{p} d_i}{\sigma^{2p}}$, and $c_3 = \exp(\frac{1}{\sigma^2} \sum_{i=1}^{p} d_i)$.
Now let's evaluate the first derivative of the denominator of Eq. (3.80).

$$
\begin{aligned}
&((\mathrm{K}_{\max})^p \cdot c_3^{\frac{1}{\mathrm{K}_{\max}}})' \\
&= p(\mathrm{K}_{\max})^{p-1} c_3^{\frac{1}{\mathrm{K}_{\max}}} + (\mathrm{K}_{\max})^p \frac{c_3^{\frac{1}{\mathrm{K}_{\max}}} \log c_3}{-(\mathrm{K}_{\max})^2} \\
&= p(\mathrm{K}_{\max})^{p-1} c_3^{\frac{1}{\mathrm{K}_{\max}}} - (\mathrm{K}_{\max})^{p-2} c_3^{\frac{1}{\mathrm{K}_{\max}}} \log c_3 \\
&= (\mathrm{K}_{\max})^{p-2} c_3^{\frac{1}{\mathrm{K}_{\max}}} (p\,\mathrm{K}_{\max} - \log c_3) \\
&= (\mathrm{K}_{\max})^{p-2} c_3^{\frac{1}{\mathrm{K}_{\max}}} (p\,\mathrm{K}_{\max} - \frac{1}{\sigma^2} \sum_{i=1}^{p} d_i)
\end{aligned}
$$

Since $d_1 > d_2 > \cdots > d_p > \sigma^2 \mathrm{K}_{\max}$,

$$
p\,\mathrm{K}_{\max} - \frac{1}{\sigma^2} \sum_{i=1}^{p} d_i < 0 \tag{3.80}
$$

This implies the denominator of Eq. (3.80) is a decreasing function, and therefore, $LR(\mathrm{K}_{\max})$ is a increasing function with respect to $\mathrm{K}_{\max}$.

(b) $p \to p+1$ as $\mathrm{K}_{\max}$ decreases

The $\mathrm{LR}(\mathrm{K}_{\max})$ is a continuous function since $\lambda_{p+1} = d_{p+1}$ at the moment that $\sigma^2 \mathrm{K}_{\max} = d_{p+1}$ and there is no discontinuity of λ_i. Therefore, $\mathrm{LR}(\mathrm{K}_{\max})$ is an increasing function in this case.

4. $d_1 > \sigma^2 K_{\max}$ and $K_{\max} < \frac{\sum_{i=1}^{c} d_i}{c - \sum_{\bar{N}+1}^{N}(d_i - 1)}$

$$\hat{\mathbf{R}}_{\text{CN}} = \mathbf{\Phi}\,\text{diag}(\boldsymbol{\lambda}^*)\mathbf{\Phi}^H \tag{3.81}$$

where

$$\boldsymbol{\lambda}^\star = [\frac{\sigma^2}{u}, \ldots, \frac{\sigma^2}{u}, d_{p+1}, \ldots, d_q, \frac{\sigma^2}{uK_{\max}}, \ldots, \frac{\sigma^2}{uK_{\max}}] \tag{3.82}$$

p, q, and $\bar{N}$ are the vector of the eigenvalues of the estimate, the largest indices so that $d_p > \frac{\sigma^2}{u}$, $d_q > \frac{\sigma^2}{u\,\mathrm{K}_{\max}}$, and $d_{\bar{N}} \geq \sigma^2$, respectively.

Before we prove the increasing property of $\mathrm{LR}(\mathrm{K}_{\max})$, we show u decreases as $\mathrm{K}_{\max}$ increases. u is the optimal solution of the optimization problem. In this case, $u^\star$ is obtained by making the first derivative of the cost function 0. Let u_1 and u_2 be the optimal solutions for $\mathrm{K}_{\max 1}$ and $\mathrm{K}_{\max 2}$, respectively. Then, $\sum_{i=1}^{N} G_i'(u_1) = 0$ for $\mathrm{K}_{\max 1}$. Since $\frac{1}{d_i} \leq u_1 \leq \frac{1}{\mathrm{K}_{\max 1}}$ in this case, for $\mathrm{K}_{\max 2} < \mathrm{K}_{\max 1}$, the value of $G_i'(u_1)$ decreases for $d_i \leq 1$. $G_i'(u)$ also decreases for $d_i > 1$ and $u \leq \frac{1}{\mathrm{K}_{\max} d_i}$ and remain same for $d_i > 1$ and $\frac{1}{\mathrm{K}_{\max} d_i} < u$. Therefore, $\sum_{i=1}^{N} G_i'(u_1) < 0$ for $\mathrm{K}_{\max 2}$. Finally, since $\sum_{i=1}^{N} G_i'(u_2)$ must be zero for $\mathrm{K}_{\max 2}$, it is obvious that $u_1 < u_2$. This shows that u decreases as $\mathrm{K}_{\max}$ increases.

Now we show the increasing property of $\mathrm{LR}(\mathrm{K}_{\max})$.

(a) within the range where p and q remain same

In this case, We show $\mathrm{LR}(u)$ is a decreasing function of u and an increasing function of $\mathrm{K}_{\max}$ for each of u and $\mathrm{K}_{\max}$.

i. Proof of $\mathrm{LR}(u)$ is a decreasing function.

$$\begin{aligned}
&\mathrm{LR}(u) \\
&= \frac{\prod_{i=1}^{p} \frac{ud_i}{\sigma^2} \cdot \prod_{i=q+1}^{\bar{N}} \frac{\mathrm{K}_{\max} ud_i}{\sigma^2} \cdot e^N}{\exp(\sum_{i=1}^{p} \frac{ud_i}{\sigma^2} + \sum_{i=p+1}^{q} 1 + \sum_{i=q+1}^{N} \frac{\mathrm{K}_{\max} ud_i}{\sigma^2})} \\
&= \frac{u^p \prod_{i=1}^{p} \frac{d_i}{\sigma^2} \cdot u^{N-q} \prod_{i=q+1}^{\bar{N}} \frac{\mathrm{K}_{\max} d_i}{\sigma^2} \cdot e^N}{\exp(u(\sum_{i=1}^{p} \frac{d_i}{\sigma^2} + \sum_{i=q+1}^{N} \frac{\mathrm{K}_{\max} d_i}{\sigma^2}) + q - p)} \\
&= \frac{c_1 u^{N-q+p}}{\exp(c_2 u + c_3)} \\
&= c_4 \frac{u^{N-q+p}}{c_5^u}
\end{aligned}$$

where $c_1 = \prod_{i=1}^{p} \frac{d_i}{\sigma^2} \cdot \prod_{i=q+1}^{\bar{N}} \frac{\mathrm{K}_{\max} d_i}{\sigma^2} \cdot e^N$, $c_2 = \sum_{i=1}^{p} \frac{d_i}{\sigma^2} + \sum_{i=q+1}^{N} \frac{\mathrm{K}_{\max} d_i}{\sigma^2}$, $c_3 = q - p$,

$c_4 = \frac{c_1}{e^{c_3}}$, and $c_5 = e^{c_2}$. The first derivative of Eq. (3.83) is obtained by

$$\begin{aligned}
&\mathrm{LR}'(u) \\
&= (N-q+p)u^{N-q+p-1}c_5^{-u} \\
&\quad - u^{N-q+p}\log c_5 \cdot c_5^{-u} \\
&= u^{N-q+p-1}c_5^{-u}(N-q+p-u\log c_5) \\
&= u^{N-q+p-1}c_5^{-u}(N-q+p-c_2u) \\
&= u^{N-q+p-1}c_5^{-u}(N-q+p \\
&\quad - u(\sum_{i=1}^{p}\frac{d_i}{\sigma^2} + \sum_{i=q+1}^{N}\frac{\mathrm{K}_{\max} d_i}{\sigma^2}))
\end{aligned}$$

Since $\frac{\sigma^2}{u} \leq d_p$,

$$\begin{aligned}
&N-q+p-u(\sum_{i=1}^{p}\frac{d_i}{\sigma^2} + \sum_{i=q+1}^{N}\frac{\mathrm{K}_{\max} d_i}{\sigma^2}) \\
&\leq N-q+p-u(\frac{p}{u} + \frac{N-q}{u}\cdot \mathrm{K}_{\max}) \\
&= N-q-\mathrm{K}_{\max}(N-q)
\end{aligned}$$

Since $\mathrm{K}_{\max} > 1$, $\mathrm{LR}'(u) < 0$ which implies $\mathrm{LR}(u)$ is a decreasing function with respect to u.

ii. Proof of $\mathrm{LR}(\mathrm{K}_{\max})$ is an increasing function.

$$\begin{aligned}
&\mathrm{LR}(\mathrm{K}_{\max}) \\
&= \frac{\prod_{i=1}^{p}\frac{ud_i}{\sigma^2}\cdot\prod_{i=q+1}^{\bar{N}}\frac{\mathrm{K}_{\max} ud_i}{\sigma^2}\cdot e^N}{\exp(\sum_{i=1}^{p}\frac{ud_i}{\sigma^2}+\sum_{i=p+1}^{q}1} \\
&\quad \overline{+\sum_{i=q+1}^{N}\frac{\mathrm{K}_{\max} ud_i}{\sigma^2})} \\
&= \frac{c_1 \mathrm{K}_{\max}{}^{N-q}}{\exp(c_2\mathrm{K}_{\max} + c_3)} \\
&= c_4\frac{\mathrm{K}_{\max}{}^{N-q}}{c_5^{\mathrm{K}_{\max}}}
\end{aligned}$$

where $c_1 = \prod_{i=1}^{p}\frac{ud_i}{\sigma^2}\cdot\prod_{i=q+1}^{\bar{N}}\frac{ud_i}{\sigma^2}\cdot e^N$, $c_2 = \sum_{i=q+1}^{N}\frac{ud_i}{\sigma^2}$, $c_3 = \sum_{i=1}^{p}\frac{ud_i}{\sigma^2}+q-p$, $c_4 = \frac{c_1}{e^{c_3}}$, and $c_5 = e^{c_2}$. The first derivative is

$$
\begin{aligned}
&\mathrm{LR}'(\mathrm{K}_{\max}) \\
&= (N-q)\,\mathrm{K}_{\max}{}^{N-q-1}\, c_5^{-\,\mathrm{K}_{\max}} \\
&\quad -\ \mathrm{K}_{\max}{}^{N-q} \log c_5 \cdot c_5^{-\,\mathrm{K}_{\max}} \qquad (3.83) \\
&= \mathrm{K}_{\max}{}^{N-q-1} \\
&\quad \times c_5^{-\,\mathrm{K}_{\max}} (N-q-\mathrm{K}_{\max} \log c_5) \qquad (3.84) \\
&= \mathrm{K}_{\max}{}^{N-q+p-1} \\
&\quad \times c_5^{-u} (N-q-c_2\,\mathrm{K}_{\max}) \qquad (3.85) \\
&= \mathrm{K}_{\max}{}^{N-q+p-1} \\
&\quad \times c_5^{-u} (N-q-\mathrm{K}_{\max} \sum_{i=q+1}^{N} \frac{ud_i}{\sigma^2}) \qquad (3.86)
\end{aligned}
$$

Since $\frac{\sigma^2}{u\,\mathrm{K}_{\max}} \leq d_{q+1}$,

$$
\begin{aligned}
&N-q-\mathrm{K}_{\max} \sum_{i=q+1}^{N} \frac{ud_i}{\sigma^2} \\
&\quad \geq\ N-q-\mathrm{K}_{\max}(\frac{N-q}{\mathrm{K}_{\max}}) = 0 \qquad (3.87)
\end{aligned}
$$

Therefore, $\mathrm{LR}'(\mathrm{K}_{\max}) \geq 0$ and $\mathrm{LR}(\mathrm{K}_{\max})$ is an increasing function with respect to $\mathrm{K}_{\max}$.

These two proofs show that $\mathrm{LR}(u, \mathrm{K}_{\max})$ is an increasing function with respect to $\mathrm{K}_{\max}$.

(b) p and q changes as $\mathrm{K}_{\max}$ decreases

The $\mathrm{LR}(u, \mathrm{K}_{\max})$ is a continuous function, and therefore, $\mathrm{LR}(u, \mathrm{K}_{\max})$ is an increasing function in this case.

□

Lemma 5 formally proves that the there exist only one optimal condition number and therefore we can find the optimal condition number numerically. The algorithm of finding the optimal condition number is shown in Algorithm 3. We first set the initial condition number as the ML condition number obtained by [25]. Then we increase or decrease the condition number to the direction where the LR value decreases. Reducing the stepsize as the direction is reversed, we find the optimal condition number as precisely as we want.

Algorithm 3 The proposed algorithm to select condition number via EL

1: Obtain the ML solution of the condition number $K_{\max_{\text{ML}}}$ by the method in [25] and set the initial value of $K_{\max} = K_{\max_{\text{ML}}}$
2: Set the initial step, $\Delta = K_{\max}/100$
3: Evaluate $\text{LR}(K_{\max} - \Delta)$, $\text{LR}(K_{\max})$, $\text{LR}(K_{\max} + \Delta)$

- if $|\,\text{LR}(K_{\max_{\text{ML}}} + \Delta) - \text{LR}_0\,| < |\,\text{LR}(K_{\max_{\text{ML}}}) - \text{LR}_0\,|$
  $\rightarrow$ increase $K_{\max}$ by Δ until it does not hold.
  $\rightarrow$ then $\Delta = -\Delta/10$
- elseif $|\,\text{LR}(K_{\max_{\text{ML}}} + \Delta) - \text{LR}_0\,| > |\,\text{LR}(K_{\max_{\text{ML}}}) - \text{LR}_0\,|$
  $\rightarrow$ decrease $K_{\max}$ by Δ until it does not hold.
  $\rightarrow$ then $\Delta = -\Delta/10$

4: Repeat Step 3 until $\Delta < 0.0001$.

3.3 Experimental Validation

3.3.1 Experimental setup

In this section, we compare the proposed methods with alternative covariance estimation algorithms and parameter estimation algorithms. Two data sets are used in the experiments: 1) a radar covariance simulation model and 2) the KASSPER dataset [8].

First, we consider a radar system with an N-element uniform linear array for the simulation model. The overall covariance which is composed of jammer and additive white noise can be modeled by

$$\mathbf{R}(n, m) = \sum_{i=1}^{J} \sigma_i^2 \operatorname{sinc}[0.5\beta_i(n-m)\phi_i] e^{j(n-m)\phi_i} + \sigma_a^2 \delta(n, m) \tag{3.88}$$

where $n, m \in \{1, \ldots, N\}$, J is the number of jammers, σ_i^2 is the power associated with the ith jammer, ϕ_i is the jammer phase angle with respect to the antenna phase center, β_i is the fractional bandwidth, σ_a^2 is the actual power level of the white disturbance term, and $\delta(n, m)$ has the value of 1 only when $n = m$ and 0 otherwise. This simulation model has been widely and very successfully used in previous literature [24, 25, 74, 33] for performance analysis.

Data from the L-band data set of KASSPER program is the other data set used in our experiments. Note that the KASSPER data set exhibits two desirable characteristics: 1) the low-rank structure of clutter and 2) the true covariance matrices for each range bin have been made available. These two characteristics facilitate comparisons via powerful figures of merit. The L-band data set consists of a data cube of 1000 range bins corresponding to the returns from a single coherent processing interval from 11 channels and 32 pulses. Therefore, the dimension of observations (or the spatio-temporal product) N is $11 \times 32 = 352$. Other key parameters are detailed in Table 2.1.

We measure the normalized signal to interference and noise ratio (SINR). The normalized SINR measure is commonly used in the radar literature and given by

$$\eta = \frac{|\mathbf{s}^H \hat{\mathbf{R}}^{-1}\mathbf{s}|^2}{|\mathbf{s}^H \hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{s}||\mathbf{s}^H \mathbf{R}^{-1}\mathbf{s}|} \tag{3.89}$$

where $\mathbf{s}$ is the spatio-temporal steering vector, $\hat{\mathbf{R}}$ is the data-dependent estimate of $\mathbf{R}$, and $\mathbf{R}$ is the true covariance matrix. It is easily seen that $0 < \eta < 1$ and $\eta = 1$ if and only if $\hat{\mathbf{R}} = \mathbf{R}$. The SINR is plotted in decibels in all our experiments, that is, $\text{SINR(dB)} = 10 \log_{10} \eta$. Therefore, $\text{SINR(dB)} \leq 0$. For the KASSPER data set, since the steering vector is a function of both azimuthal angle and Doppler frequency, we obtain plots as a function of one variable (azimuthal angle or Doppler) by marginalizing over the other variable. We evaluate and compare different covariance estimation techniques and parameter selection methods in the following three subsections:

- **Sample Covariance Matrix:** The sample covariance matrix is given by $\mathbf{S} = \frac{1}{K}\mathbf{Z}\mathbf{Z}^H$. It is well known that $\mathbf{S}$ is the unconstrained ML estimator under Gaussian disturbance statistics. We refer to this as SMI.

- **Fast Maximum Likelihood:** The fast maximum likelihood (FML) [24] uses the structural constraint of the covariance matrix. The FML method just involves the eigenvalue decomposition of the sample covariance and perturbing eigenvalues to conform to the structure. The FML also can be considered as the RCML estimator with the rank which is the greatest index i satisfying $\lambda_i > \sigma^2$ where λ_i's are the eigenvalues of the sample covariance in descending order. Therefore, a rank can be considered as an output of the FML. The FML's success in radar STAP is widely known [11].

- **Rank Constrained ML Estimators:** The RCML estimator with the rank or the rank and the noise level obtained by the proposed methods using the expected likelihood approach. The rank is obtained by the EL approach in the case of the imperfect rank constraint and both of the rank and the noise level are obtained by the EL approach in the case of imperfect rank and noise power constraints. We refer to these as RCML_{EL}.

- **Chen *et al.* Rank Selection Method:** Chen *et al.* [84] proposed a statistical procedure for detecting the multiplicity of the smallest eigenvalue of the structured covariance matrix using statistical selection theory. The rank can be estimated from their methods using pre-calculated parameters. We refer to this method as $\text{RCML}_{\text{Chen}}$.

- **AIC:** Akaike [77] proposed the information theoretic criteria for model selection. The AIC selects the model that best fits the data for given a set of observations and a family of models, that is, a parameterized family of probability densities. Wax and Kailath [79] proposed the method to determine the number of signals from the observed data based on the AIC. We compare Wax and Kailath's method.

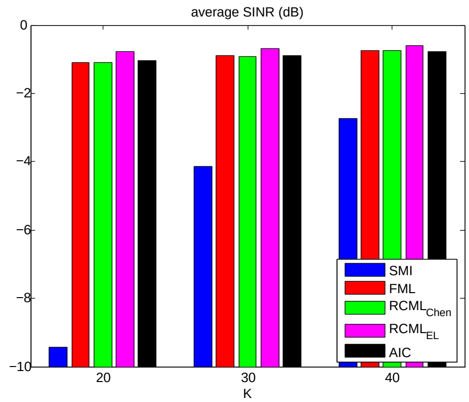


Figure 3.4: Normalized SINR in dB versus number of training samples K ($N = 20$) for the simulation model.

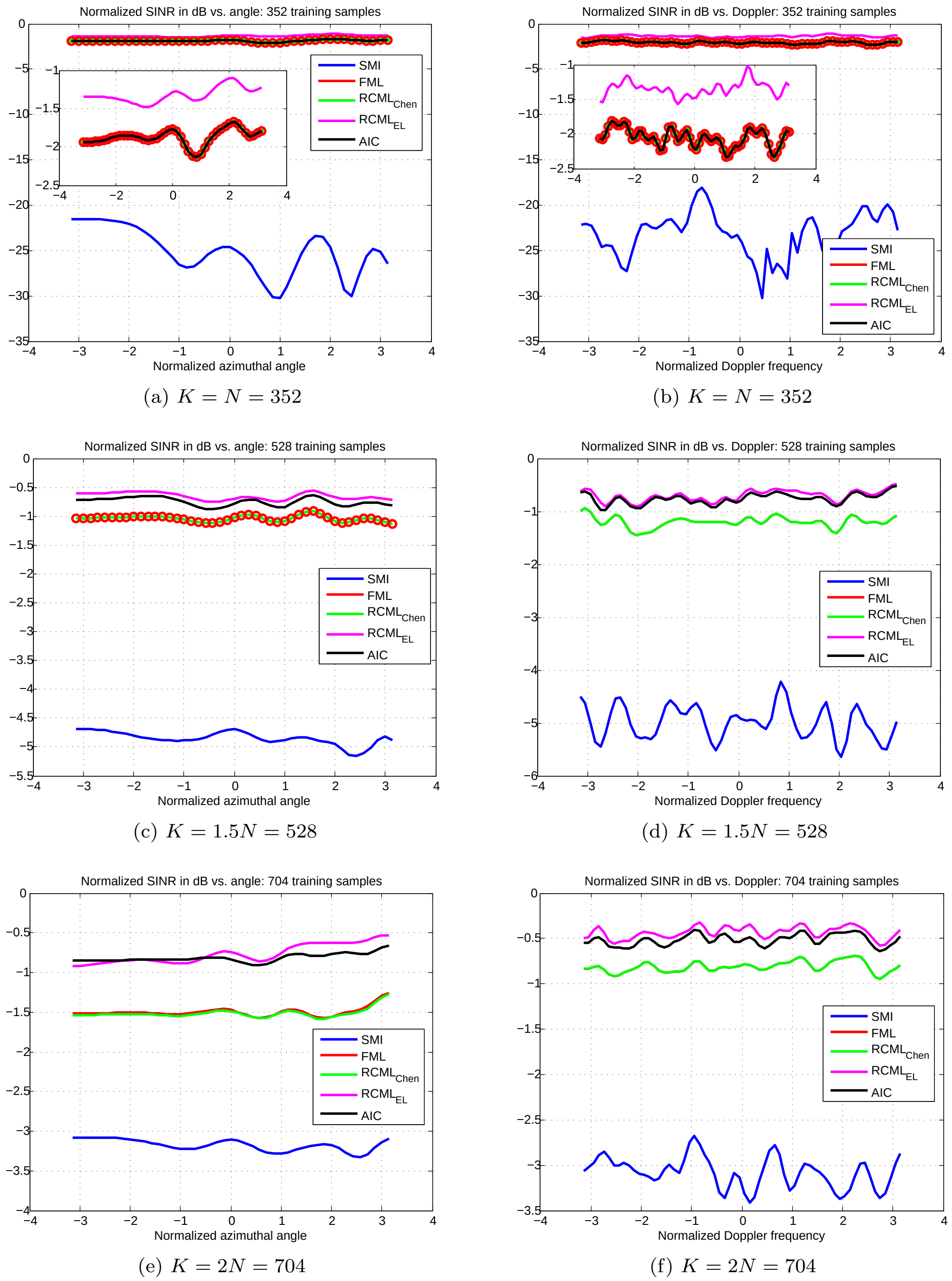
- **Condition number constrained ML estimators:** The maximum likelihood estimation method of the covariance matrix with a condition number [25] proposed by Aubry *et al.* is considered for evaluating the performance with three different condition numbers. 1) CNCML : the condition number obtained by the proposed method in [25], 2) $\mathrm{CNCML}_{\mathrm{EL}}$: the condition number obtained by the expected likelihood approach, and 3) $\mathrm{CNCML}_{\mathrm{true}}$: the true condition number.

3.3.2 Imperfect rank constraint

First, we compare the rank estimation method proposed in Section 3.2.1 with alternative algorithms including SMI, FML, AIC, and Chen's algorithm. We plot the normalized SINR (in dB) versus the number of training samples, 20, 30, and 40 in Fig. 3.4 for the simulation model. The SINR values are obtained by averaging SINR values from 500 Monte Carlo trials. It is shown that the SINR values increases monotonically as K increases. The $\mathrm{RCML}_{\mathrm{EL}}$exhibits the best performance in all training regimes. Particularly, the difference between $\mathrm{RCML}_{\mathrm{EL}}$and other methods increases when training samples are limited.

Figure 3.5 shows the normalized SINR values for various number of training samples for the KASSPER data set. We plot the averaged SINR values in decibel over either azimuth angle or Doppler frequency domain. The left and right column show the results for angle and Doppler, respectively. Similarly to the results for the simulation model, $\mathrm{RCML}_{\mathrm{EL}}$outperforms all the other compared methods in all training regime. This implies that the rank obtained by the EL approach is more accurate and closer to the rank predicted by Brennan rule ($M + P - 1 = 42$) than any other methods.

Realistic case of contaminated observations: In practice, homogeneous training samples are hard to obtain and the received signals are often corrupted by target information. Therefore, it is meaningful to compare the performance for nonhomogeneous observation to investigate which



(a) $K = N = 352$ (b) $K = N = 352$

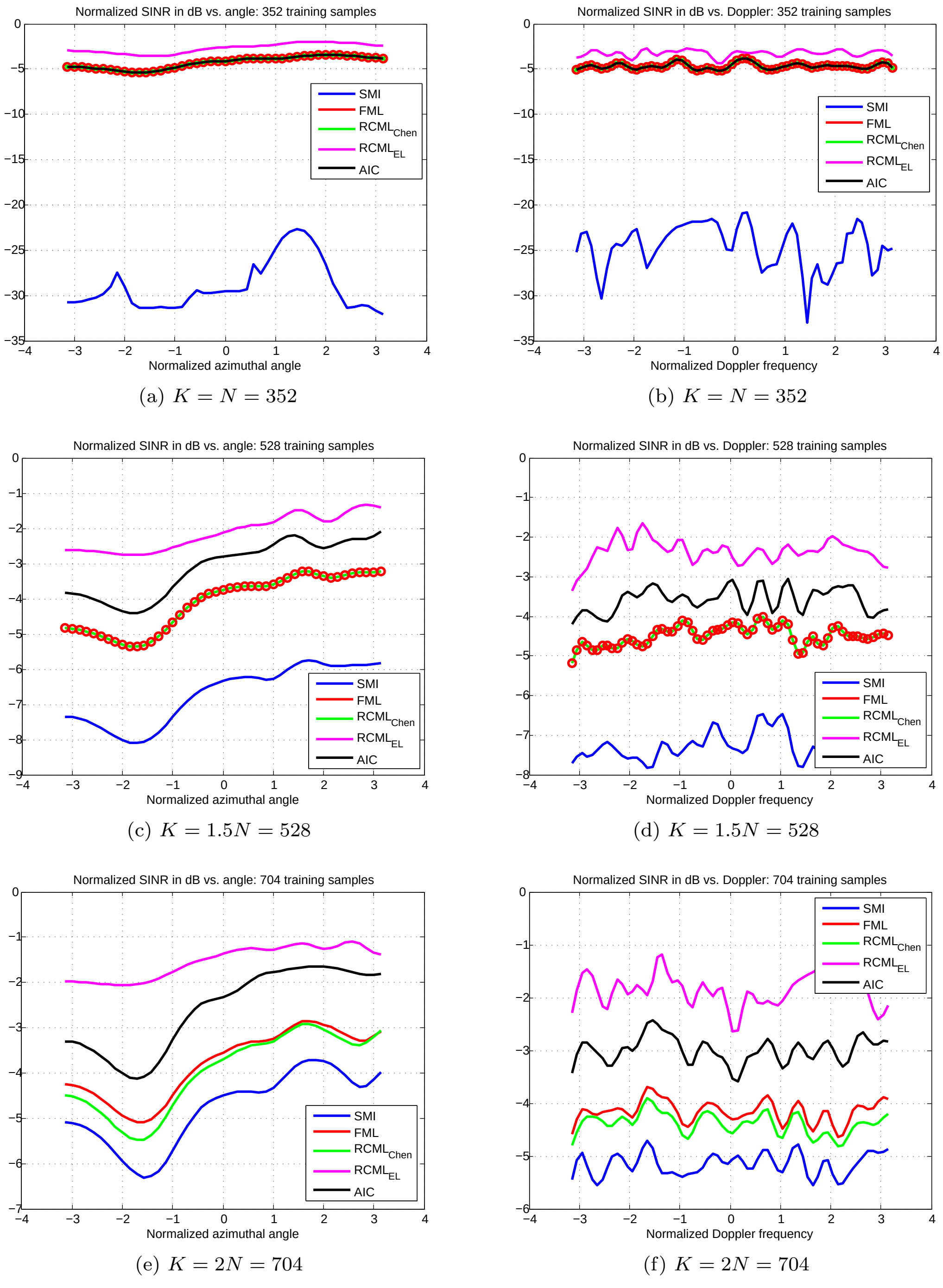
(c) $K = 1.5N = 528$ (d) $K = 1.5N = 528$

(e) $K = 2N = 704$ (f) $K = 2N = 704$

Figure 3.5: Normalized SINR versus azimuthal angle and Doppler frequency for the KASSPER data set.

algorithm indeed works well and is robust in practice. In this case, the received signal is given by

$$\mathbf{z} = \alpha \mathbf{s} + \mathbf{d} \tag{3.90}$$



(a) $K = N = 352$

(b) $K = N = 352$

(c) $K = 1.5N = 528$

(d) $K = 1.5N = 528$

(e) $K = 2N = 704$

(f) $K = 2N = 704$

Figure 3.6: Normalized SINR versus azimuthal angle and Doppler frequency for the KASSPER data set.

where $\mathbf{s}$ and $\mathbf{d}$ are the deterministic steering vector and stochastic disturbance vector, respectively. Figure 3.6 shows the normalized SINR values when a half of the training samples contain $\mathbf{s}$ with $\alpha = 50$. The gap between $\mathrm{RCML_{EL}}$and the others is bigger than that in Figure 3.5. In particular, AIC shows compatible performance in homogeneous cases though, Figure 3.6e and Figure 3.6f show that the

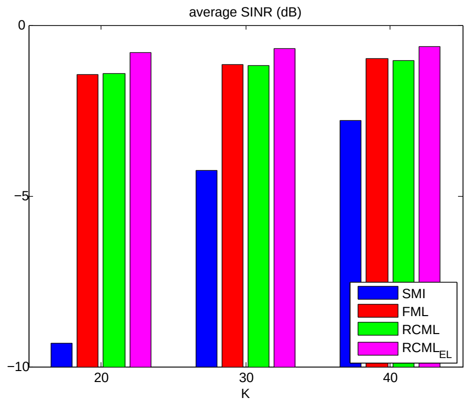


Figure 3.7: Normalized SINR in dB versus number of training samples K ($N = 20$) for the simulation model.

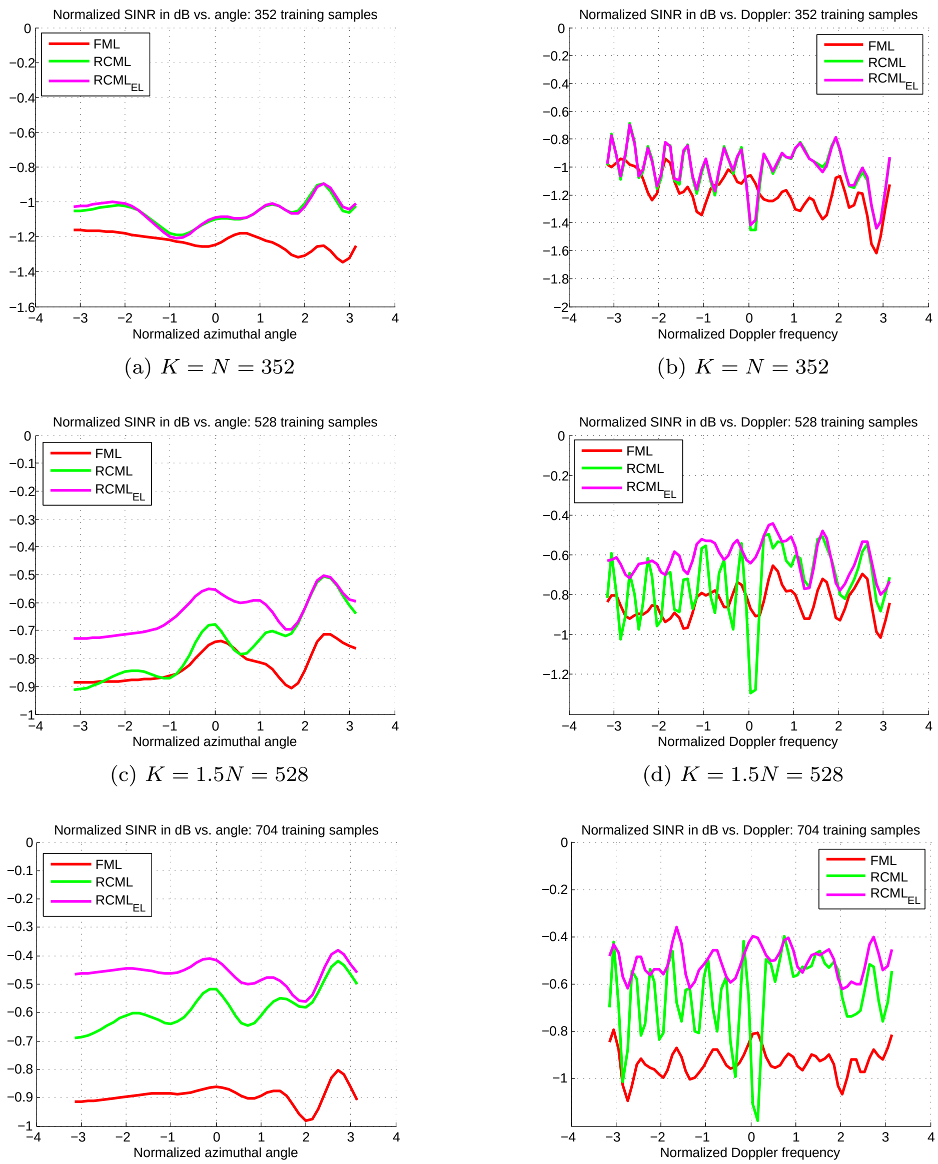
difference between AIC and RCML$_{\mathrm{EL}}$is larger for non-homogeneous case. Therefore the results show remarkably that RCML$_{\mathrm{EL}}$still excels under target contamination or heterogeneous training where other techniques face severe degradation in performance.

3.3.3 Imperfect rank and noise power constraints

In this section, we show experimental results for estimation of both a rank and a noise power via the expected likelihood approach, which is proposed in Section 3.2.2. In this case, we assume that both the rank and the noise power are unknown and to be estimated for both the simulation model and the KASSPER data set. Since the previous works such as AIC and Chen's algorithm are for only estimating the rank and can not be extended to estimate both the rank and the noise power, we compare the proposed EL method with the sample covariance, FML, and the RCML estimator with a prior knowledge of the rank. For the RCML estimator, we employ the number of jammers ($r = 5$) and the Brennan rule ($r = 42$) as the clutter rank for the simulation model and the KASSPER data set, respectively. In addition, since the FML method requires a prior knowledge of the noise power, we calculate and use the maximum likelihood estimate of the noise power for a rank given by a prior knowledge for the FML.

Figure 3.7 shows the performance of various estimators in the sense of the normalized SINR values for the simulation model. Similarly to the case of only rank estimation, the proposed method show the best performance in all training regimes.

Figure 3.8 shows the performance of the methods in the normalized SINR for the KASSPER data set. The proposed method is comparable with or slightly better than the RCML estimator using the rank by Brennan rule. This means that the proposed method estimates both the rank and the noise power adaptively from training samples whereas the rank by Brennan rule is fixed regardless of the training samples.



(a) $K = N = 352$

(b) $K = N = 352$

(c) $K = 1.5N = 528$

(d) $K = 1.5N = 528$

(e) $K = 2N = 704$

(f) $K = 2N = 704$

Figure 3.8: Normalized SINR versus azimuthal angle and Doppler frequency for the KASSPER data set.

3.3.4 Imperfect condition number constraint

Now we show experimental results for the condition number estimation method proposed in Section 3.2.3. We compare the proposed method, denoted by $\mathrm{CNCML}_{\mathrm{EL}}$, with four different covariance estimation methods, the sample covariance matrix (SMI), FML, CNCML , and $\mathrm{CNCML}_{\mathrm{true}}$.

Tables 3.1 and 3.2 show the normalized SINR values for the simulation model. We analyze five different scenarios with different parameters of the simulated covariance model given by Eq. (3.88). We use the same parameters as those used in [25] to evaluate the performances and they are shown

σ^2	K	SMI	FML	CNCML	CNCML_{EL}	$\text{CNCML}_{\text{true}}$
-5	20	-9.3785	-0.5195	-0.5212	**-0.4822**	-0.5200
	30	-4.2579	**-0.4242**	-0.4257	-0.4256	-0.4250
	40	-2.7424	**-0.3460**	-0.3476	-0.3476	-0.3468
0	20	-9.3196	-0.5511	-0.5521	**-0.5141**	-0.5515
	30	-4.2276	**-0.4202**	-0.4221	-0.4220	-0.4210
	40	-2.7649	**-0.3513**	-0.3530	-0.3528	-0.3521
5	20	-9.0922	-0.5269	-0.5279	**-0.4875**	-0.5272
	30	-4.2172	**-0.4348**	-0.4364	-0.4362	-0.4357
	40	-2.7300	**-0.3484**	-0.3503	-0.3505	-0.3493
10	20	-9.3511	-0.5355	-0.5305	**-0.4998**	-0.5360
	30	-4.1955	**-0.4164**	-0.4180	-0.4175	-0.4175
	40	-2.7491	**-0.3501**	-0.3515	-0.3518	-0.3509

(a)

σ^2	K	SMI	FML	CNCML	CNCML_{EL}	$\text{CNCML}_{\text{true}}$
-5	20	-9.3069	-1.7371	**-1.7322**	-1.7358	-1.7350
	30	-4.1795	-1.2399	-1.2388	**-1.2347**	-1.2397
	40	-2.7535	-0.9496	-0.9492	**-0.9456**	-0.9493
0	20	-9.1354	-1.6944	**-1.6928**	-1.7027	-1.6940
	30	-4.2345	-1.2986	-1.2987	**-1.2955**	-1.2990
	40	-2.7545	-1.0041	-1.0043	**-1.0023**	-1.0046
5	20	-9.2524	-1.3976	-1.4016	**-1.3244**	-1.4000
	30	-4.2309	-1.0737	-1.0784	**-1.0666**	-1.0762
	40	-2.7523	-0.8848	-0.8876	**-0.8818**	-0.8866
10	20	-9.3660	-1.2567	-1.2569	**-1.2115**	-1.2570
	30	-4.3013	-0.9526	-0.9545	**-0.9450**	-0.9537
	40	-2.7350	-0.7171	-0.7197	**-0.7139**	-0.7186

(b)

σ^2	K	SMI	FML	CNCML	CNCML_{EL}	$\text{CNCML}_{\text{true}}$
-5	20	-9.3702	-0.5340	-0.5349	**-0.4925**	-0.5340
	30	-4.2791	**-0.4302**	-0.4316	-0.4315	-0.4308
	40	-2.7856	**-0.3493**	-0.3510	-0.3509	-0.3501
0	20	-9.2898	-0.5485	-0.5501	**-0.5104**	-0.5491
	30	-4.2648	**-0.4202**	-0.4219	-0.4220	-0.4209
	40	-2.7274	**-0.3604**	-0.3621	-0.3621	-0.3611
5	20	-9.0582	-0.5318	-0.5328	**-0.4899**	-0.5322
	30	-4.1548	**-0.4142**	-0.4155	-0.4152	-0.4149
	40	-2.7655	**-0.3515**	-0.3531	-0.3533	-0.3521
10	20	-9.3632	-0.5352	-0.5363	**-0.4974**	-0.5360
	30	-4.2728	**-0.4328**	-0.4348	-0.4349	-0.4337
	40	-2.7577	**-0.3538**	-0.3554	-0.3547	-0.3547

(c)

Table 3.1: Normalized SINR for various values of parameters for the simulation model.

in Table 5.2(c).

For the narrowband scenarios ($B_f = 0$) in Table 5.1(a) and Table 5.1(c), CNCML_{EL} outperforms the alternatives for the limited training regime and FML is the best in other training regimes. Note

σ^2	K	SMI	FML	CNCML	CNCML_{EL}	$\text{CNCML}_{\text{true}}$
-5	20	-9.0316	-1.7161	**-1.7131**	-1.7634	-1.7150
	30	-4.1465	-1.1704	-1.1691	**-1.1659**	-1.1693
	40	-2.7727	-0.9390	-0.9384	**-0.9351**	-0.9387
0	20	-9.2091	-1.6706	-1.6701	**-1.6674**	-1.6706
	30	-4.2004	-1.2681	-1.2682	**-1.2633**	-1.2682
	40	-2.7423	-1.0102	-1.0117	**-1.1009**	-1.0106
5	20	-9.3538	-1.3980	-1.4028	**-1.3216**	-1.4004
	30	-4.2203	-1.0869	-1.0910	**-1.0785**	-1.0889
	40	-2.7079	-0.8694	-0.8721	**-0.8666**	-0.8713
10	20	-9.221	-1.2446	-1.2455	**-1.1982**	-1.2452
	30	-4.2116	-0.9428	-0.9460	**-0.9382**	-0.9444
	40	-2.7563	-0.7235	-0.7264	**-0.7226**	-0.7253

(a)

σ^2	K	SMI	FML	CNCML	CNCML_{EL}	$\text{CNCML}_{\text{true}}$
-5	20	-9.2679	-1.1593	-1.1616	**-1.1150**	-1.1610
	30	-4.2234	-0.9262	-0.9286	**-0.9242**	-0.9278
	40	-2.8271	**-0.7705**	-0.7729	-0.7712	-0.7723
0	20	-9.2934	-0.9052	-0.9051	**-0.8422**	-0.9059
	30	-4.1617	-0.6909	-0.6920	**-0.6862**	-0.6924
	40	-2.7387	-0.5711	-0.5724	**-0.5676**	-0.5725
5	20	-9.4154	-0.8398	-0.8334	**-0.7909**	-0.8399
	30	-4.2284	-0.6273	-0.6231	**-0.6070**	-0.6278
	40	-2.7208	-0.5034	-0.5022	**-0.4945**	-0.5046
10	20	-9.1447	-0.7388	-0.7225	**-0.6815**	-0.7392
	30	-4.2046	-0.5931	-0.5803	**-0.5535**	-0.5931
	40	-2.7241	-0.4821	-0.4738	**-0.4576**	-0.4827

(b)

	J	σ_J^2	ϕ	B_f
(a)	1	30	20°	0
(b)	1	30	20°	0.3
(c)	3	[30 30 30]	[20° 40° 60°]	[0 0 0]
(d)	3	[30 30 30]	[20° 40° 60°]	[0.3 0.3 0.3]
(e)	3	[10 20 30]	[20° 40° 60°]	[0.2 0 0.3]

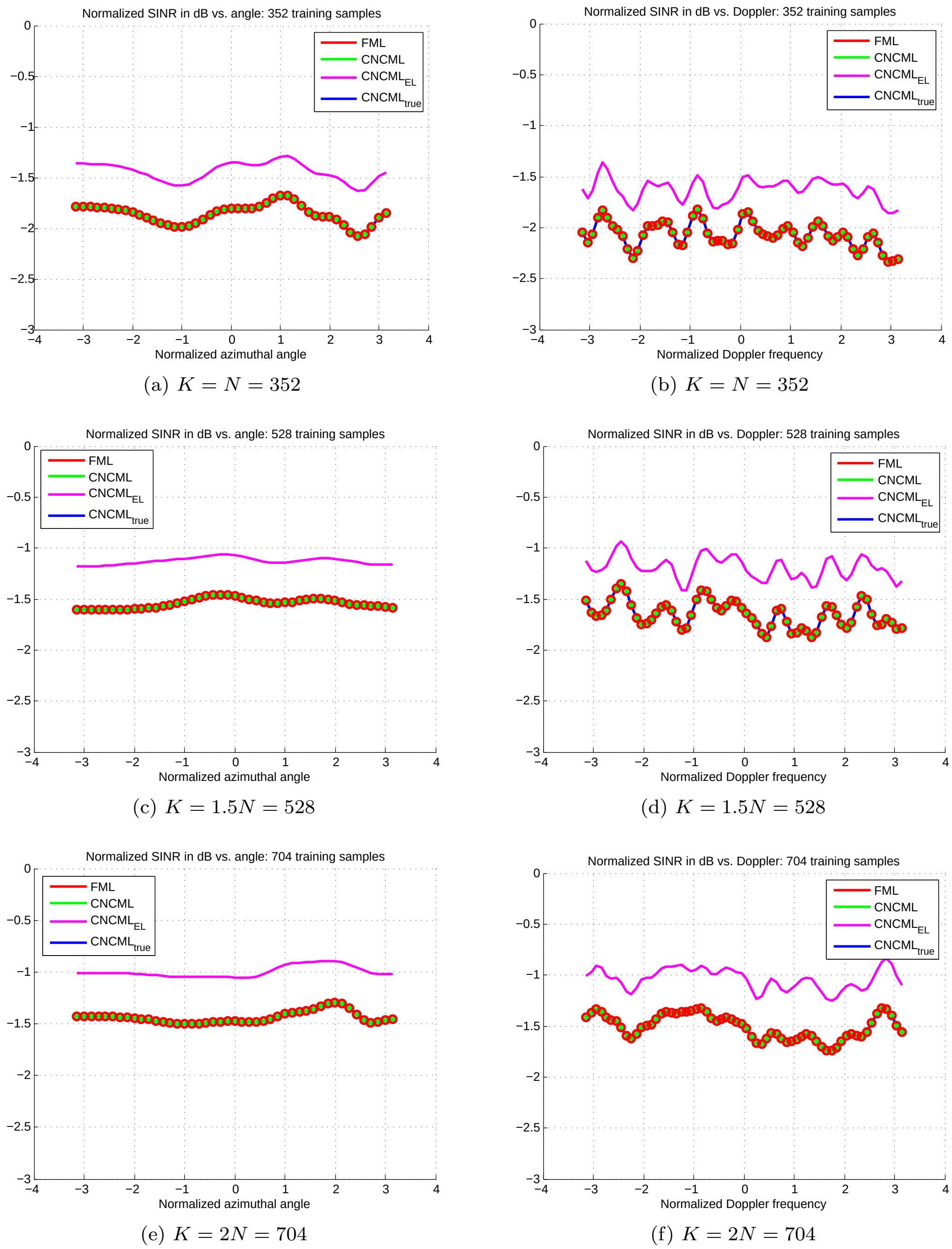
(c)

Table 3.2: Normalized SINR for various values of parameters for the simulation model.

that the gap between CNCML_{EL} and FML (at most 0.002) is much smaller than that of the limited training regime (at least 0.3).

On the other hand, for the wideband scenarios in Tables 5.1(b), 5.2(a), and 5.2(b), CNCML_{EL} shows the best performance in most cases including $\text{CNCML}_{\text{true}}$ using the true condition number.

The experimental results for the KASSPER data set are shown in Figure 3.9. We do not plot the sample covariance matrix to clarify the difference among the estimators. In every case, FML,



(a) $K = N = 352$

(b) $K = N = 352$

(c) $K = 1.5N = 528$

(d) $K = 1.5N = 528$

(e) $K = 2N = 704$

(f) $K = 2N = 704$

Figure 3.9: Normalized SINR versus azimuthal angle and Doppler frequency for the KASSPER data set. (a) and (b) for $K = N = 352$, (c) and (d) for $K = 1.5N = 528$, and (e) and (f) for $K = 2N = 704$

CNCML, and $\text{CNCML}_{\text{true}}$ are very close to one another and CNCML_{EL} is the best estimator. Note that CNCML_{EL} is based on the same algorithm as CNCML and $\text{CNCML}_{\text{true}}$ and differs from them in the point that CNCML_{EL} uses a different condition number which is estimated by the expected likelihood approach. Again, this shows that the expected likelihood criterion is really useful and powerful to estimate parameters which is imperfectly known and leads to an adaptive and robust covariance estimator.

3.4 Conclusion

We propose robust covariance estimation algorithms which automatically determine the optimal values of practical constraints via the expected likelihood criterion for radar STAP in this chapter. Three different cases of practical constraints which is exploited in recent works including the rank constrained ML estimation and the condition number constrained ML estimation are investigated. Significant analytic results and proofs are derived for each case. Uniqueness of the optimal values of the rank constraint and the condition number constraint is formally proved and a closed form solution of the noise level is derived. Experimental results show that the estimators with the constraints obtained by the expected likelihood outperform previous works which uses constraints obtained by other parameter estimation methods including the maximum likelihood solution of the constraints.

Bibliography

[1] J. Ward, "Space-Time Adpative Processing for Airborne Radar," Tech. Rep., MIT Lincoln Laboratory, December 1994.

[2] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nizberg, "A CFAR Adaptive Matched Filter Detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, January 1992.

[3] B. Widrow, P. E. Mantey, L. J. Griffths, and B. B. Goode, "Adaptive Antenna Systems," in *Proceedings of the IEEE*, December 1967, vol. 55, pp. 2143–2159.

[4] S. P. Applebaum, "Adaptive arrays," Tech. Rep., Syracuse University Research Corporation, August 1966.

[5] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid Convergence Rate in Adaptive Arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-10, no. 6, pp. 853–863, 1974.

[6] F. Gini and M. Rangaswamy, *Knowledge Based Radar Detection, Tracking and Classification*, Wiley-Interscience publication, 2008.

[7] B. Himed and W. L. Melvin, "Analyzing Space-Time Adaptive Processors Using Measured Data," in *Conference Record of The Thirty-First Asilomar Conference on Signals, Systems and Computers*, November 1997, vol. 1, pp. 930–935.

[8] J. Bergin and P. Techau, "High-fidelity site specific radar simulation: Kassper'02 workshop datacube," Tech. Rep. ISL-SCRD-TR-02-105, DAPRA Technical Report, May 2002.

[9] M. Rangaswamy, J. H. Michels, and B. Himed, "Statistical analysis of the non-homogeneity detector for STAP applications," *Digital Signal Processing*, vol. 14, no. 3, pp. 253–267, May 2004.

[10] M. Rangaswamy, "Statistical Analysis of the Nonhomogeneity Detector for Non-Gaussian Interference Backgrounds," *IEEE Trans. Signal Processing*, vol. 53, no. 6, pp. 2101–2111, June 2005.

[11] M. Rangaswamy, F. C. Lin, and K. R. Gerlach, "Robust adaptive signal processing methods for heterogeneous radar clutter scenarios," *Signal Processing*, vol. 84, no. 9, pp. 1653–1665, September 2004.

[12] H. Wang and L. Cai, "A Localized Adaptive MTD Processor," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 3, pp. 532–539, May 1991.

[13] J. R. Roman, M. Rangaswamy, D. W. Davis, Q. Zhang, B. Himed, and J. H. Michels, "Parametric Adaptive Matched Filter for Airborne Radar Applications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 677–692, April 2000.

[14] J. S. Goldstein, L. L. Scharf, and I. S. Reed, "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2943–2959, November 1998.

[15] T. K. Sarkar and N. Sangruji, "An Adaptive Nulling System for a Narrow-Band Signal with a Look-Direction Constraint Utilizing the Conjugate Gradient Method," *IEEE Trans. Antennas Propagat.*, vol. 37, no. 7, pp. 940–944, July 1989.

[16] R. S. Adve, T. B. Hale, and M. C. Wicks, "Practical joint domain localised adaptive processing in homogeneous and nonhomogeneous environments. Part 2: Nonhomogeneous environments," in *IEE Proc. Radar, Sonar and Navig.*, April 2000, vol. 147, pp. 66–74.

[17] R. Nitzberg, "Application of Maximum Likelihood Estimation of Persymmetric Covariance Matrices to Adaptive Processing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-16, no. 1, pp. 124–127, January 1980.

[18] D. R. Fuhrmann, "Application of Toeplitz Covariance Estimation to Adaptive Beamforming and Detection," *IEEE Trans. Signal Processing*, vol. 39, no. 10, pp. 2194–2198, October 1991.

[19] H. Li, P. Stoica, and J. Li, "Computationally Efficient Maximum Likelihood Estimation of Structured Covariance Matrices," *IEEE Trans. Signal Processing*, vol. 47, no. 5, pp. 1314–1323, May 1999.

[20] Y. I. Abramovich, D. A. Gray, A. Y. Gorokhov, and N. K. Spencer, "Positive-Definite Toeplitz Completion in DOA Estimation for Nonuniform Linear Antenna Arrays-Part I: Fully Augmentable Arrays," *IEEE Trans. Signal Processing*, vol. 46, no. 9, pp. 2458–2471, September 1998.

[21] E. Conte, M. Lops, and G. Ricci, "Adaptive Detection Schemes in Compound-Gaussian Clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 4, pp. 1058–1069, October 1998.

[22] P. Wang, H. Li, and B. Himed, "A Simplified Parametric GLRT for STAP Detection," in *IEEE Radar Conference*, May 2009, pp. 1–5.

[23] A. L. Kraay and A. B. Baggeroer, "A Physically Constrained Maximum-Likelihood Method for Snapshot-Deficient Adaptive Array Processing," *IEEE Trans. Signal Processing*, vol. 55, no. 8, pp. 4048–4063, August 2007.

[24] M. Steiner and K. Gerlach, "Fast Converging Adaptive Processors for a Structured Covariance Matrix," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 4, pp. 1115–1126, October 2000.

[25] A. Aubry, A. De Maio, L. Pallotta, and A. Farina, "Maximum Likelihood Estimation of a Structured Covariance Matrix with a Condition Number Constraint," *IEEE Trans. Signal Processing*, vol. 60, no. 6, pp. 3004–3021, June 2012.

[26] A. De Maio and A. Farina, "Adaptive Radar Detection: A Bayesian Approach," in *International Radar Symposium*, May 2006, pp. 1–4.

[27] A. De Maio, A. Farina, and G. Foglia, "Adaptive Radar Detection: A Bayesian Approach," in *IEEE Radar Conference*, April 2007, pp. 624–629.

[28] O. Besson, J. Y. Tourneret, and S. Bidon, "Knowledge-Aided Bayesian Detection in Heterogeneous Environments," *IEEE Signal Processing Lett.*, vol. 14, no. 5, pp. 355–358, May 2007.

[29] J. R. Guerci and E. J. Baranoski, "Knowledge-Aided Adaptive Radar at DARPA: An overview," *IEEE Signal Processing Mag.*, vol. 23, no. 1, pp. 41–50, January 2006.

[30] P. Wang, H. Li, and B. Himed, "A Bayesian Parametric Test for Multichannel Adaptive Signal Detection in Nonhomogeneous Environments," *IEEE Signal Processing Lett.*, vol. 17, no. 4, pp. 351–354, April 2010.

[31] P. R. Gurram and N. A. Goodman, "Spectral-Domain Covariance Estimation with A Priori Knowledge," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 3, pp. 1010–1020, July 2006.

[32] W. L. Melvin and G. A. Showman, "An Approach to Knowledge-Aided Covariance Estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 3, pp. 1021–1042, July 2006.

[33] A. De Maio, S. De Nicola, L. Landi, and A. Farina, "Knowledge-Aided Covariance Matrix Estimation: a MAXDET Approach," *IET Radar, Sonar and Navigation*, vol. 3, no. 4, pp. 341–356, August 2009.

[34] A. De Maio, G. Foglia, A. Farina, and M. Piezzo, "Estimation of the Covariance Matrix Based on Multiple A-Priori Models," in *IEEE Radar Conference*, May 2010, pp. 1025–1029.

[35] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero, "Shrinkage Algorithms for MMSE Covariance Estimation," *IEEE Trans. Signal Processing*, vol. 58, no. 10, pp. 5016–5029, October 2010.

[36] O. Ledoit and M. Wolf, "Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection," *Journal of Empirical Finance*, vol. 10, no. 5, pp. 603–621, December 2003.

[37] P. Stoica, J. Li, X. Zhu, and J. R. Guerci, "On Using A-Priori Knowledge in Space-Time Adaptive Processing," *IEEE Trans. Signal Processing*, vol. 56, no. 6, pp. 2598–2602, June 2008.

[38] J. H. Won, J. Lim, S. J. Kim, and B. Rajaratnam, "Maximum Likelihood Covariance Estimation with a Condition Number Constraint," Tech. Rep., Department of Statistics, Stanford University, August 2009.

[39] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2nd edition, 2004.

[40] M. C. Wicks, M. Rangaswamy, R. Adve, and T. B. Hale, "Space-Time Adaptive Processing: A knowledge-based perspective for airborne radar," *IEEE Signal Processing Mag.*, vol. 23, no. 1, pp. 51–65, January 2006.

[41] J. P. Hoffbeck and D. A. Landgrebe, "Covariance Matrix Estimation and Classification with Limited Training Data," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, no. 7, pp. 763–767, July 1996.

[42] A. Haimovich, "The Eigencanceler: Adaptive Radar by Eigenanalysis Methods," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 532–542, April 1996.

[43] I. Kirsteins and D. Tufts, "Adaptive Detection Using Low Rank Approximation to a Data Matrix," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 30, no. 1, pp. 55–67, January 1994.

[44] R. A. Monzingo and T. Miller, *Introduction to Adaptive Arrays*, SciTech Publishing, 1st edition, 2004.

[45] S. D. Blunt, K. R. Gerlach, and M. Rangaswamy, "The Enhanced FRACTA Algorithm With Knowledge-Aided Covariance Estimation," in *Proc. Third IEEE Workshop on Sensor Array and Multichannel Processing*, April 2004, pp. 638–642.

[46] E. J. Kelly, "An Adaptive Detection Algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 1, pp. 115–127, March 1986.

[47] J. R. Guerci, *Space-Time Adpative Processing for Radar*, Artech House Publishers, July 2003.

[48] R. Klemm, *Principles of space-time adaptive processing*, IEE Publishing, April 2002.

[49] G. T. Capraro, A. Farina, H. Griffiths, and M. C. Wicks, "Knowledge-Based Radar Signal and Data Processing: A tutorial overview," *IEEE Signal Processing Mag.*, vol. 23, no. 1, pp. 18–29, January 2006.

[50] S. Haykin, "Cognitive Radar: A way of the future," *IEEE Signal Processing Mag.*, vol. 23, no. 1, pp. 30–40, January 2006.

[51] S. Miranda, C. Baker, K. Woodbridge, and H. Griffiths, "Knowledge-Based Resource Management for Multifunction Radar: A look at scheduling and task prioritization," *IEEE Signal Processing Mag.*, vol. 23, no. 1, pp. 66–76, January 2006.

[52] A. Aubry, A. De Maio, L. Pallota, and A. Farina, "Radar Detection of Distributed Targets in Homogeneous Interference Whose Inverse Covariance Structure is Defined via Unitary Invariant Functions," *IEEE Trans. Signal Processing*, vol. 61, no. 20, pp. 4949–4961, October 2013.

[53] P. Chen, W. L. Melvin, and M. C. Wicks, "Screening among Multivariate Normal Data," *Journal of Multivariate Analysis*, vol. 69, no. 1, pp. 10–29, April 1999.

[54] A. Aubry, A. De Maio, L. Pallotta, and A. Farina, "Covariance Matrix Estimation via Geometric Barycenters and its Application to Radar Training Data Selection," *IET Radar, Sonar and Navigation*, vol. 7, no. 6, pp. 600–614, July 2013.

[55] V. Monga and M. Rangaswamy, "Rank Constrained ML Estimation of Structured Covariance Matrices with Applications in Radar Target Detection," in *IEEE Radar Conference*, May 2012, pp. 475–480.

[56] B. Kang, V. Monga, and M. Rangaswamy, "On the Practical Merits of Rank constrained ML Estimator of Structured Covariance Matrices," in *IEEE Radar Conference*, May 2013, pp. 1–6.

[57] J. P. Burg, D. G. Luenberger, and D. L. Wenger, "Estimation of Structured Covariance Matrices," *Proceedings of the IEEE*, vol. 70, no. 9, pp. 963–974, September 1982.

[58] P. Stoica and T. Söderström, "On reparameterization of loss functions used in estimation and the invariance principle," *Signal Processing*, vol. 17, no. 4, pp. 383–387, August 1989.

[59] S. S. Al-Homidan, "SQP Algorithms for Solving Toeplitz Matrix Approximation Problem," *Numerical Linear Algebra with Applications*, vol. 9, no. 8, pp. 619–627, 2002.

[60] L. Davis, R. Evans, and E. Polak, "Maximum Likelihood Estimation of Positive Definite Hermitian Toeplitz Matrices Using Outer Approximations," in *Statistical Signal and Array Processing, 1998. Proceedings., Ninth IEEE SP Workshop on*, September 1998, pp. 49–52.

[61] A. K. Shaw, S. Pokala, and R. Kumaresan, "Toeplitz and Hankel Matrix Approximation using Structured Approach," in *Acoustics, Speech, and Signal Processing, 1998. ICASSP-98., 1998 International Conference on*, May 1998, vol. 4, pp. 2349–2352.

[62] T. J. Suffridge and T. L. Hayden, "Approximation by a Hermitian Positive Semidefinite Toeplitz Matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 14, no. 3, pp. 721–734, July 1993.

[63] M. I. Miller and D. L. Snyder, "The Role of Likelihood and Entropy in Incomplete-Data Problems: Application to Estimating Point-Process Intensities and Toeplitz Constrained Covariances," *Proceedings of the IEEE*, vol. 75, no. 7, pp. 892–907, July 1987.

[64] R. J. A. Little and D. B. Rubin, *Statistical Analysis with Missing Data*, A John Wiley and Sons, 2nd edition, 2002.

[65] P. Forster and G. Vezzosi, "Optimal Toeplitzification with a Given Rank Constraint," in *Acoustics, Speech, and Signal Processing, 1989. ICASSP-89., 1989 International Conference on*, May 1989, vol. 4, pp. 2783–2786.

[66] M. Jansson and B. Ottersten, "Structured Covariance Matrix Estimation: A Parametric Approach," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00., 2000 IEEE International Conference on*, June 2000, vol. 5, pp. 3172–3175.

[67] M. E. Tipping and C. M. Bishop, "Probabilistic Principal Component Analysis," *Journal of the Royal Statistical Society*, vol. 61, no. 3, pp. 611–622, October 1999.

[68] P. Stoica and M. Jansson, "On Maximum Likelihood Estimation in Factor Analysis - An Algebraic Derivation," *Signal Processing*, vol. 89, no. 6, pp. 1260–1262, June 2009.

[69] N. A. Goodman and J. M. Stiles, "On Clutter Rank Observed by Arbitrary Arrays," *IEEE Trans. Signal Processing*, vol. 55, no. 1, pp. 178–186, January 2007.

[70] B. Kang, V. Monga, and M. Rangaswamy, "Rank-Constrained Maximum Likelihood Estimation of Structured Covariance Matrices," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 1, pp. 501–516, January 2014.

[71] M. Skolnik, *Radar Handbook*, The McGraw-Hill Professional, 3rd edition, 2008.

[72] D. M. Wilkes and M. H. Hayes, "Iterated Toeplitz Approximation of Covariance Matrices," in *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on*, April 1988, vol. 3, pp. 1663–1666.

[73] C. Eckart and G. Young, "The Approximation of One Matrix by Another of Lower Rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, September 1936.

[74] L. Pallotta, A. Aubry, A. De Maio, and A. Farina, "Estimation of a Structured Covariance Matrix with a Condition Number Constraint for Radar Applications," in *IEEE Radar Conference*, May 2012, pp. 778–783.

[75] E. Conte, M. Lops, and G. Ricci, "Asymptotically Optimum Radar Detection in Compound-Gaussian Clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 2, pp. 617–625, April 1995.

[76] A. Aubry, A. De Maio, and V. Carotenuto, "Optimality Claims for the FML Covariance Estimator with respect to Two Matrix Norms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 3, pp. 2055–2057, July 2013.

[77] H. Akaike, "A New Look at the Statistical Model Indentification," *IEEE Trans. Auto. Contr.*, vol. AC-19, no. 6, pp. 716–723, December 1974.

[78] J. Rissanen, "Modeling by the Shortest Data Description," *Automatica*, vol. 14, no. 5, pp. 465–471, September 1978.

[79] M. Wax and T. Kailath, "Detection of Signals by Information Theoretic Criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 2, pp. 387–392, April 1985.

[80] Y. Q. Yin and P. R. Krishnaiah, "On Some Nonparametric Methods for Detection of the Number of signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, no. 11, pp. 1533–1538, November 1987.

[81] D. W. Tufts and A. A. Shah, "Rank Determination in Time-Series Analysis," in *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 International Conference on*, April 1994, vol. 4, pp. IV21–IV24.

[82] Y. I. Abramovich, N. K. Spencer, and A. Y. Gorokhov, "Modified GLRT and AMF Framework for Adaptive Detectors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 3, pp. 1017–1051, July 2007.

[83] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lamber W function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, 1996.

[84] P. Chen, M. C. Wicks, and R. S. Adve, "Development of a statistical procedure for detecting the number of signals in a radar measurement," *IEE Proceedings – Radar, Sonar and Navigation*, vol. 148, no. 4, pp. 219–226, August 2001.

1.

1. Report Type

Final Report

Primary Contact E-mail
Contact email if there is a problem with the report.

vmonga@engr.psu.edu

Primary Contact Phone Number
Contact phone number if there is a problem with the report

8148631267

Organization / Institution name

The Pennsylvania State University

Grant/Contract Title
The full title of the funded effort.

Robust, Adaptive Radar Detection and Estimation

Grant/Contract Number
AFOSR assigned control number. It must begin with "FA9550" or "F49620" or "FA2386".

FA9550-12-1-0333

Principal Investigator Name
The full name of the principal investigator on the grant or contract.

Vishal Monga

Program Manager
The AFOSR Program Manager currently assigned to the award

Arje Nachman

Reporting Period Start Date

07/15/2012

Reporting Period End Date

07/14/2015

Abstract

This work introduces estimation of disturbance covariance matrices for radar STAP. In particular, we first exploit physically inspired constraints, both the structure of the disturbance covariance and importantly the knowledge of the clutter rank to yield a new rank constrained maximum likelihood (RCML) estimator of clutter/disturbance covariance. We demonstrate that the rank-constrained estimation problem can in fact be cast in the framework of a tractable convex optimization problem, and derive closed form expressions for the estimated covariance matrix. On top of that, this work also introduces a new computationally efficient covariance estimator which jointly enforces a Toeplitz structure and a rank constraint. Our proposed solution focuses on a computationally efficient approximation and involves a cascade of two closed form solutions, the RCML estimator and the rank preserving Toeplitz approximation. Finally, we develop robust estimators that can adapt to imperfect knowledge of physical constraints using an expected likelihood (EL) approach. We analyze covariance estimation algorithms under three cases of imperfect constraints: only rank constraint, both rank and noise power constraints, and condition number constraint. Extensive performance evaluation on both simulated and KASSPER data conrms the proposed estimators excel previous works.

Distribution Statement
This is block 12 on the SF298 form.

Distribution A - Approved for Public Release

Explanation for Distribution Statement

If this is not approved for public release, please provide a short explanation. E.g., contains proprietary information.

SF298 Form

Please attach your SF298 form. A blank SF298 can be found here. Please do not password protect or secure the PDF The maximum file size for an SF298 is 50MB.

AFD-070820-035.pdf

Upload the Report Document. File must be a PDF. Please do not password protect or secure the PDF . The maximum file size for the Report Document is 50MB.

FinalReport_2015.pdf

Upload a Report Document, if any. The maximum file size for the Report Document is 50MB.

Archival Publications (published) during reporting period:

1. V. Monga and M. Rangaswamy, "Rank Constrained ML Estimation of Structured Covariance Matrices with Applications in Radar Target Detection," 2012 IEEE Radar Conference

2. B. Kang, V. Monga, and M. Rangaswamy, "On the Practical Merits of Rank Constrained ML Estimator of Structured Covariance Matrices," 2013 IEEE Radar Conference

3. B. Kang, V. Monga, and M. Rangaswamy, "Efficient Approximation of Structured Covariance under Joint Toeplitz and Rank Constraints," IEEE Asilomar Conference on Signal, Systems and Computers, November 2013

4. B. Kang, V. Monga, and M. Rangaswamy, "Constrained ML Estimation of Structured Covariance Matrices with Applications in Radar STAP," Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), The fifth IEEE International Workshop, December 2013

5. B. Kang, V. Monga, and M. Rangaswamy, "Rank Constrained ML Estimation of Structured Covariance Matrices," IEEE Transactions on Aerospace and Electronic Systems, January 2014

6. B. Kang, V. Monga, and M. Rangaswamy, "Estimation of Structured Covariance Matrices for Radar STAP under Practical Constraints," 2014 IEEE Radar Conference, The Best Student Paper Award

7. B. Kang, V. Monga, and M. Rangaswamy, "Computationally Efficient Toeplitz Approximation of Structured Covariance under a Rank Constraint," IEEE Transactions on Aerospace and Electronic Systems, January 2015

8. B. Kang, "Student Research Highlight: Estimation of Structured Covariance Matrices for Radar STAP," IEEE Aerospace and Electronic Systems Magazine, February 2015

9. B. Kang, V. Monga, M. Rangaswamy, and Y. Abramovich, "Automatic Rank Estimation for Practical STAP Covariance Estimation via an Expected Likelihood Approach," 2015 IEEE Radar Conference

10. B. Kang, V. Monga, M. Rangaswamy, and Y. Abramovich, "Robust Covariance Estimation under Imperfect Constraints using Expected Likelihood Approach," to be submitted to IEEE Transactions on Aerospace and Electronic Systems

Changes in research objectives (if any):

None

Change in AFOSR Program Manager, if any:

None

Extensions granted or milestones slipped, if any:

None

AFOSR LRIR Number

LRIR Title

Reporting Period

Laboratory Task Manager

Program Officer

Research Objectives

Technical Summary

Funding Summary by Cost Category (by FY, $K)

	Starting FY	FY+1	FY+2
Salary			
Equipment/Facilities			
Supplies			
Total			

Report Document

Report Document - Text Analysis

Report Document - Text Analysis

Appendix Documents

2. Thank You

E-mail user

Jul 14, 2015 15:59:08 Success: Email Sent to: vmonga@engr.psu.edu